



## 特許協力条約に基づく国際出願願書

2F00089-PCT

原本 (出願用) - 印刷日時 2000年08月22日 (22.08.2000) 火曜日 16時29分30秒

0	受理官庁記入欄	
0-1	国際出願番号.	
0-2	国際出願日	
0-3	(受付印)	
0-4	様式-PCT/R0/101 この特許協力条約に基づく国際出願願書は、 右記によって作成された。	PCT-EASY Version 2.91 (updated 01.07.2000)
0-5	申立て 出願人は、この国際出願が特許協力条約に従って処理されることを請求する。	
0-6	出願人によって指定された受理官庁	日本国特許庁 (R0/JP)
0-7	出願人又は代理人の書類記号	2F00089-PCT
I	発明の名称	音声符号化装置及び音声符号化方法
II	出願人	出願人である (applicant only)
II-1	この欄に記載した者は	米国を除くすべての指定国 (all designated States except US)
II-2	右の指定国についての出願人である。	松下電器産業株式会社 MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD. 571-8501 日本国 大阪府 門真市 大字門真1006番地 1006, Oaza Kadoma, Kadoma-shi, Osaka 571-8501 Japan
II-4ja	名称	
II-4en	Name	
II-5ja	あて名:	
II-5en	Address:	
II-6	国籍 (国名)	日本国 JP
II-7	住所 (国名)	日本国 JP
II-8	電話番号	06-6908-1473
II-9	ファクシミリ番号	06-6909-0053
III-1	その他の出願人又は発明者	出願人及び発明者である (applicant and inventor)
III-1-1	この欄に記載した者は	米国のみ (US only)
III-1-2	右の指定国についての出願人である。	安永 和敏 YASUNAGA, Kazutoshi 216-0015 日本国 神奈川県 川崎市 宮前区菅生3-33-17-305 3-33-17-305, Sugao, Miyamae-ku, Kawasaki-shi, Kanagawa 216-0015 Japan
III-1-4ja	氏名 (姓名)	
III-1-4en	Name (LAST, First)	
III-1-5ja	あて名:	
III-1-5en	Address:	
III-1-6	国籍 (国名)	日本国 JP
III-1-7	住所 (国名)	日本国 JP



---

**THIS PAGE BLANK (USPTO)**

## 特許協力条約に基づく国際出願願書

2F00089-PCT

原本（出願用） - 印刷日時 2000年08月22日 (22.08.2000) 火曜日 16時29分30秒

III-2 III-2-1	その他の出願人又は発明者 この欄に記載した者は	出願人及び発明者である (applicant and inventor)
III-2-2	右の指定国についての出願人である。	米国のみ (US only)
III-2-4ja	氏名 (姓名)	森井 利幸
III-2-4en	Name (LAST, First)	MORI, Kazutoshi
III-2-5ja	あて名:	215-0015 日本国 神奈川県 川崎市
III-2-5en	Address:	麻生区虹ヶ丘2-3-7-501 2-3-7-501, Nijigaoka, Asao-ku, Kawasaki-shi, Kanagawa 215-0015
III-2-6	国籍 (国名)	日本国 JP
III-2-7	住所 (国名)	日本国 JP
IV-1	代理人又は共通の代表者、通知のあて名	代理人 (agent)
IV-1-1ja	下記の者は国際機関において右記のごとく出願人のために行動する。	鷺田 公一
IV-1-1en	氏名 (姓名)	WASHIDA, Kimihito
IV-1-2ja	Name (LAST, First)	206-0034 日本国 東京都 多摩市
IV-1-2en	あて名:	鶴牧1丁目24-1 新都市センタービル5階
IV-1-2en	Address:	5th Floor, Shintoshicenter Bldg., 24-1, Tsurumaki 1-chome, Tama-shi, Tokyo 206-0034
IV-1-3	電話番号	Japan 042-338-4600
IV-1-4	ファクシミリ番号	042-338-4605
V	国の指定	
V-1	広域特許 (他の種類の保護又は取扱いを求める場合には括弧内に記載する。)	AP: GH GM KE LS MW MZ SD SL SZ TZ UG ZW 及びハラレプロトコルと特許協力条約の締約国である他の国
		EA: AM AZ BY KG KZ MD RU TJ TM 及びユーラシア特許条約と特許協力条約の締約国である他の国
		EP: AT BE CH&LI CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE 及びヨーロッパ特許条約と特許協力条約の締約国である他の国
		OA: BF BJ CF CG CI CM GA GN GW ML MR NE SN TD TG 及びアフリカ知的所有権機構と特許協力条約の締約国である他の国
V-2	国内特許 (他の種類の保護又は取扱いを求める場合には括弧内に記載する。)	AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH&LI CN CR CU CZ DE DK DM DZ EE ES FI GB GD GE GH GM HR HU ID IL IN IS KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT TZ UA UG US UZ VN YU ZA ZW

---

**THIS PAGE BLANK (USPTO)**



## 特許協力条約に基づく国際出願願書

2F00089-PCT

原本（出願用） - 印刷日時 2000年08月22日（22.08.2000）火曜日 16時29分30秒

V-5	指定の確認の宣言 出願人は、上記の指定に加えて、規則4.9(b)の規定に基づき、特許協力条約のもとで認められる他の全ての国の指定を行う。ただし、V-6欄に示した国の指定を除く。出願人は、これらの追加される指定が確認を条件としていること、並びに優先日から15月が経過する前にその確認がなされない指定は、この期間の経過時に、出願人によって取り下げられたものとみなされることを宣言する。		
V-6	指定の確認から除かれる国	なし (NONE)	
VI-1	先の国内出願に基づく優先権主張		
VI-1-1	先の出願日	1999年08月23日 (23.08.1999)	
VI-1-2	先の出願番号	特願平11-235050	
VI-1-3	国名	日本国 JP	
VI-2	先の国内出願に基づく優先権主張		
VI-2-1	先の出願日	1999年08月24日 (24.08.1999)	
VI-2-2	先の出願番号	特願平11-236728	
VI-2-3	国名	日本国 JP	
VI-3	先の国内出願に基づく優先権主張		
VI-3-1	先の出願日	1999年09月02日 (02.09.1999)	
VI-3-2	先の出願番号	特願平11-248363	
VI-3-3	国名	日本国 JP	
VI-4	優先権証明書送付の請求 上記の先の出願のうち、右記の番号のものについては、出願書類の認証謄本を作成し国際事務局へ送付することを、受理官庁に対して請求している。	VI-1, VI-2, VI-3	
VII-1	特定された国際調査機関 (ISA)	日本国特許庁 (ISA/JP)	
VIII	照合欄	用紙の枚数	添付された電子データ
VIII-1	願書	4	-
VIII-2	明細書	47	-
VIII-3	請求の範囲	3	-
VIII-4	要約	1	2f00089-pct.txt
VIII-5	図面	14	-
VIII-7	合計	69	


---

THIS PAGE BLANK (USPTO)

## 特許協力条約に基づく国際出願願書

2F00089-PCT

原本（出願用） - 印刷日時 2000年08月22日（22.08.2000）火曜日 16時29分30秒

	添付書類	添付	添付された電子データ
VIII-8	手数料計算用紙	✓	-
VIII-9	別個の記名押印された委任状	✓	-
VIII-10	包括委任状の写し	✓	-
VIII-16	PCT-EASYディスク	-	フレキシブルディスク
VIII-17	その他	国際事務局の口座への振込を証明する書面	-
VIII-17	その他	納付する手数料に相当する特許印紙を貼付した書面	-
VIII-18	要約書とともに提示する図の番号	4	
VIII-19	国際出願の使用言語名:	日本語 (Japanese)	
IX-1	提出者の記名押印		
IX-1-1	氏名(姓名)	鷲田 公一	

## 受理官庁記入欄

10-1	国際出願として提出された書類の実際の受理の日	
10-2	図面:	
10-2-1	受理された	
10-2-2	不足図面がある	
10-3	国際出願として提出された書類を補完する書類又は図面であってその後期間内に提出されたものの実際の受理の日（訂正日）	
10-4	特許協力条約第11条(2)に基づく必要な補完の期間内の受理の日	
10-5	出願人により特定された国際調査機関	ISA/JP
10-6	調査手数料未払いにつき、国際調査機関に調査用写しを送付していない	

## 国際事務局記入欄

II-1	記録原本の受理の日	
------	-----------	--

---

**THIS PAGE BLANK (USPTO)**

## 明 細 書

## 音声符号化装置及び音声符号化方法

## 5 技術分野

本発明は、デジタル通信システムにおいて使用される音声符号化装置および音声符号化方法に関する。

## 背景技術

- 10 携帯電話などのデジタル移動通信の分野では、加入者の増加に対処するために低ビットレートの音声の圧縮符号化法が求められており、各研究機関において研究開発が進んでいる。

- 日本国内においては、モトローラ社が開発したビットレート11.2 kbpsのVSELPという符号化法がデジタル携帯電話用の標準符号化方式として採用され、同方式を搭載したデジタル携帯電話は1994年秋から国内において発売されている。

- また、NTT移動通信網株式会社の開発したビットレート5.6 kbpsのPSI-CELPという符号化方式が現在製品化されている。これらの方式はいずれもCELP(Code Excited Linear Prediction: M.R.Schroeder"  
20 High Quality Speech at Low Bit Rates" Proc.ICASSP'85 pp.937-940に記載されている)という方式を改良したものである。

- このCELP方式は、音声を音源情報と声道情報とに分離し、音源情報については符号帳に格納された複数の音源サンプルのインデクスによって符号化し、声道情報についてはLPC(線形予測係数)を符号化すること及び  
25 音源情報符号化の際に声道情報を加味して入力音声とを比較することを行う方法(A-b-S:Analysis by Synthesis)を採用していることに特徴がある。

このCELP方式においては、まず、入力された音声データ（入力音声）に対して自己相関分析とLPC分析を行ってLPC係数を得て、得られたLPC係数の符号化を行ってLPC符号を得る。さらに、得られたLPC符号を復号化して復号化LPC係数を得る。一方、入力音声は、LPC係数を用いた聴感

5 重み付けフィルタを用いて聴感重み付けされる。

適応符号帳と確率的符号帳に格納された音源サンプル（それぞれ適応コードベクトル（又は適応音源）、確率的コードベクトル（又は、確率的音源）と呼ぶ）のそれぞれのコードベクトルに対して、得られた復号化LPC係数によってフィルタリングを行い、2つの合成音を得る。

- 10 そして、得られた2つの合成音と、聴感重み付けされた入力音声との関係进行分析し、2つの合成音の最適値（最適ゲイン）を求め、求められた最適ゲインによって合成音をパワー調整し、それぞれの合成音を加算して総合合成音を得る。その後、得られた総合合成音と入力音声との間の符号化歪みを求める。このようにして、全ての音源サンプルに対して総合合成音と入力音声との間の符号化歪みを求め、符号化歪みが最も小さいときの音源サンプルのインデックスを
- 15 求める。

- このようにして得られたゲイン及び音源サンプルのインデックスを符号化し、これらの符号化されたゲイン及び音源サンプルをLPC符号と共に伝送路に送る。また、ゲイン符号と音源サンプルのインデックスに対応する2つの音源から実際の音源信号を作成し、それを適応符号帳に格納すると同時に古い音源サンプルを破棄する。
- 20

なお、一般的には、適応符号帳と確率的符号帳に対する音源探索は、分析区間をさらに細かく分けた区間（サブフレームと呼ばれる）で行われる。

- ゲインの符号化（ゲイン量子化）は、音源サンプルのインデックスに対応する2つの合成音を用いてゲインの量子化歪を評価するベクトル量子化（VQ）によって行われる。
- 25

このアルゴリズムにおいては、予めパラメータベクトルの代表的サンプル

(コードベクトル) が複数格納されたベクトル符号帳を作成しておく。次いで、聴感重み付けした入力音声と、適応音源及び確率的音源を聴感重み付けLPC合成したものに対して、ベクトル符号帳に格納されたゲインコードベクトルを用いて符号化歪を下記式1により計算する。

$$5 \quad E_n = \sum_{i=0}^I (X_i - g_n \times A_i - h_n \times S_i)^2$$

式1

ここで、

$E_n$  : n番のゲインコードベクトルを用いたときの符号化歪み

$X_i$  : 聴感重み付け音声

10  $A_i$  : 聴感重み付けLPC合成済み適応音源

$S_i$  : 聴感重み付けLPC合成済み確率的音源

$g_n$  : コードベクトルの要素 (適応音源側のゲイン)

$h_n$  : コードベクトルの要素 (確率的音源側のゲイン)

$n$  : コードベクトルの番号

15  $i$  : 音源データのインデクス

$I$  : サブフレーム長 (入力音声の符号化単位)

次いで、ベクトル符号帳を制御することによって各コードベクトルを用いたときの歪 $E_n$ を比較し、最も歪の小さいコードベクトルの番号をベクトルの符号とする。また、ベクトル符号帳に格納された全てのコードベクトルの中で最も歪みが小さくなるコードベクトルの番号を求め、これをベクトルの符号とする。

上記式1は一見して各n毎に多くの計算を必要とるように見えるが、予めiについての積和を計算しておけばよいので、少ない計算量でnの探索を行うことができる。

25 一方、音声復号化装置 (デコーダ) では、伝送されてきたベクトルの符号に基づいてコードベクトルを求めることによって符号化されたデータを復号化

してコードベクトルを得る。

また、上記アルゴリズムを基本として、従来よりさらなる改良がなされてきた。例えば、人間の音圧の聴覚特性が対数であることを利用し、パワを対数化

して量子化し、そのパワで正規化した2つのゲインをVQする。この方法は、

- 5 日本国PDCハーフレートコーデックの標準方式で用いられている方法である。また、ゲインパラメータのフレーム間相関を利用して符号化する方法（予測符号化）がある。この方法は、ITU-T国際標準G. 729で用いられている方法である。しかしながら、これらの改良によっても十分な性能を得ることができていない。

- 10 これまで人間の聴覚特性やフレーム間相関を利用したゲイン情報符号化法が開発され、ある程度効率の良いゲイン情報の符号化が可能になった。特に、予測量子化によって性能は大きく向上したが、その従来法では、状態としての値として以前のサブフレームの値をそのまま用いて予測量子化を行っていた。しかしながら、状態として格納される値の中には、極端に大きな（小さな）値
- 15 をとるものがあり、その値を次のサブフレームに用いると、次のサブフレームの量子化がうまくいかず、局所的異音になる場合がある。

#### 発明の開示

- 本発明の目的は、予測量子化を用いて局所的異音を生じることなく音声符号
- 20 化を行うことができるCELP型音声符号化装置及び方法を提供することである。

本発明の主題は、予測量子化において前のサブフレームでの状態値が極端に大きな値や極端に小さな値である場合に、自動的に予測係数を調整することにより、局所的異音の発生を防止することである。

25

#### 図面の簡単な説明

図1は、本発明の音声符号化装置を備えた無線通信装置の構成を示すブロッ



ク図；

図 2 は、本発明の実施の形態 1 に係る音声符号化装置の構成を示すブロック図；

図 3 は、図 2 に示す音声符号化装置におけるゲイン演算部の構成を示すブロック図；

図 4 は、図 2 に示す音声符号化装置におけるパラメータ符号化部の構成を示すブロック図；

図 5 は、本発明の実施の形態 1 に係る音声符号化装置で符号化された音声データを復号する音声復号化装置の構成を示すブロック図；

10 図 6 は、適応符号帳探索を説明するための図；

図 7 は、本発明の実施の形態 2 に係る音声符号化装置の構成を示すブロック図；

図 8 は、パルス拡散符号帳を説明するためのブロック図；

図 9 は、パルス拡散符号帳の詳細な構成の一例を示すブロック図；

15 図 10 は、パルス拡散符号帳の詳細な構成の一例を示すブロック図；

図 11 は、本発明の実施の形態 3 に係る音声符号化装置の構成を示すブロック図；

図 12 は、本発明の実施の形態 3 に係る音声符号化装置で符号化された音声データを復号する音声復号化装置の構成を示すブロック図；

20 図 13 A は、本発明の実施の形態 3 に係る音声符号化装置で用いるパルス拡散符号帳の一例を示す図；

図 13 B は、本発明の実施の形態 3 に係る音声復号化装置で用いるパルス拡散符号帳の一例を示す図；

図 14 A は、本発明の実施の形態 3 に係る音声符号化装置で用いるパルス拡散符号帳の一例を示す図；並びに

図 14 B は、本発明の実施の形態 3 に係る音声復号化装置で用いるパルス拡散符号帳の一例を示す図である。

発明を実施するための最良の形態

以下、本発明の実施の形態について、添付図面を参照して詳細に説明する。

(実施の形態 1)

- 5 図 1 は、本発明の実施の形態 1 ～ 3 に係る音声符号化装置を備えた無線通信装置の構成を示すブロック図である。

この無線通信装置において、送信側で音声が入力装置 1 1 によって電氣的アナログ信号に変換され、A/D変換器 1 2 に出力される。アナログ音声信号は、A/D変換器 1 2 によってデジタル音声信号に変換され、  
10 音声符号化部 1 3 に出力される。音声符号化部 1 3 は、デジタル音声信号に対して音声符号化処理を行い、符号化した情報を変復調部 1 4 に出力する。変復調部 1 4 は、符号化された音声信号をデジタル変調して、無線送信部 1 5 に送る。無線送信部 1 5 では、変調後の信号に所定の無線送信処理を施す。この信号は、アンテナ 1 6 を介して送信される。なお、プロセッサ 2 1 は、適宜  
15 RAM 2 2 及び ROM 2 3 に格納されたデータを用いて処理を行う。

一方、無線通信装置の受信側では、アンテナ 1 6 で受信した受信信号は、無線受信部 1 7 で所定の無線受信処理が施され、変復調部 1 4 に送られる。変復調部 1 4 では、受信信号に対して復調処理を行い、復調後の信号を音声復号化部 1 8 に出力する。音声復号化部 1 8 は、復調後の信号に復号処理を行ってデ  
20 ィジタル復号音声信号を得て、そのデジタル復号音声信号を D/A 変換器 1 9 へ出力する。D/A 変換器 1 9 は、音声復号化部 1 8 から出力されたデジタル復号音声信号をアナログ復号音声信号に変換してスピーカなどの音声出力装置 2 0 に出力する。最後に音声出力装置 2 0 が電氣的アナログ復号音声信号を復号音声に変換して出力する。

25 ここで、音声符号化部 1 3 及び音声復号化部 1 8 は、RAM 2 2 及び ROM 2 3 に格納された符号帳を用いて DSP などのプロセッサ 2 1 により動作する。また、これらの動作プログラムは、ROM 2 3 に格納されている。

図2は、本発明の実施の形態1に係るCELP型音声符号化装置の構成を示すブロック図である。この音声符号化装置は、図1に示す音声符号化部13に含まれている。なお、図2に示す適応符号帳103は図1に示すRAM22に格納されており、図2に示す確率的符号帳104は図1に示すROM23に格納されている。

図2に示す音声符号化装置においては、LPC分析部102において、入力された音声データ101に対して自己相関分析及びLPC分析を行ってLPC係数を得る。また、LPC分析部102では、得られたLPC係数の符号化を行ってLPC符号を得る。さらに、LPC分析部102では、得られたLPC符号を復号化して復号化LPC係数を得る。入力された音声データ101は、聴感重み付け部107に送られ、そこで上記LPC係数を用いた聴感重み付けフィルタを用いて聴感重み付けされる。

次に、音源作成部105において、適応符号帳103に格納された音源サンプル（適応コードベクトル又は適応音源）と確率的符号帳104に格納された音源サンプル（確率的コードベクトル又は、確率的音源）を取り出し、それぞれのコードベクトルを聴感重みLPC合成部106へ送る。さらに、聴感重みLPC合成部106において、音源作成部105で得られた2つの音源に対して、LPC分析部102で得られた復号化LPC係数によってフィルタリングを行い、2つの合成音を得る。

なお、聴感重みLPC合成部106においては、LPC係数や高域強調フィルタや長期予測係数（入力音声の長期予測分析を行うことによって得られる）を用いた聴感重み付けフィルターを併用してそれぞれの合成音に対して聴感重み付けLPC合成を行う。

聴感重みLPC合成部106は、2つの合成音をゲイン演算部108に出力する。ゲイン演算部108は、図3に示す構成を有する。ゲイン演算部108においては、聴感重みLPC合成部106で得られた2つの合成音及び聴感重み付けされた入力音声を分析部1081に送り、そこで2つの合成音と入力

音声との関係进行分析し、2つの合成音の最適値（最適ゲイン）を求める。この最適ゲインは、パワ調整部1082に出力される。

パワ調整部1082では、求められた最適ゲインによって2つの合成音をパワ調整する。パワ調整された合成音は、合成部1083に出力されて、そこで  
5 加算されて総合合成音となる。この総合合成音は、符号化歪算出部1084に出力される。符号化歪算出部1084では、得られた総合合成音と入力音声との間の符号化歪みを求める。

符号化歪算出部1084は、音源作成部105を制御して、適応符号帳103及び確率的符号帳104の全ての音源サンプルを出力させ、全ての音源サン  
10 プルに対して総合合成音と入力音声との間の符号化歪みを求め、符号化歪みが最も小さいときの音源サンプルのインデクスを求める。

次に、分析部1081は、音源サンプルのインデクス、そのインデクスに対応する2つの聴感重み付けLPC合成された音源、及び入力音声をパラメータ符号化部109に送る。

15 パラメータ符号化部109では、ゲインの符号化を行うことによってゲイン符号を得、LPC符号、音源サンプルのインデクスをまとめて伝送路へ送る。また、ゲイン符号とインデクスに対応する2つの音源から実際の音源信号を作成し、それを適応符号帳103に格納すると同時に古い音源サンプルを破棄する。なお、一般的には、適応符号帳と確率的符号帳に対する音源探索は、分析  
20 区間をさらに細かく分けた区間（サブフレームと呼ばれる）で行われる。

ここで、上記構成を有する音声符号化装置のパラメータ符号化部109のゲイン符号化の動作について説明する。図4は、本発明の音声符号化装置のパラメータ符号化部の構成を示すブロック図である。

図4において、聴感重み付け入力音声（ $X_i$ ）、聴感重み付けLPC合成済み適応音源（ $A_i$ ）、及び聴感重み付けLPC合成済み確率的音源（ $S_i$ ）が  
25 パラメータ計算部1091に送られる。パラメータ計算部1091では、符号化歪計算に必要なパラメータを計算する。パラメータ計算部1091で計算され

たパラメータは、符号化歪計算部 1092 に出力され、そこで符号化歪が計算される。この符号化歪は、比較部 1093 に出力される。比較部 1093 では、符号化歪計算部 1092 及びベクトル符号帳 1094 を制御して、得られた符号化歪から最も適当とされる符号（復号化ベクトル）を求め、この符号を基に  
5 ベクトル符号帳 1094 から得られるコードベクトルを復号化ベクトル格納部 1096 に出力し、復号化ベクトル格納部 1096 を更新する。

予測係数格納部 1095 は、予測符号化に用いる予測係数を格納する。この予測係数はパラメータ計算及び符号化歪計算に用いられるために、パラメータ計算部 1091 及び符号化歪計算部 1092 に出力される。復号化ベクトル格  
10 納部 1096 は、予測符号化のために状態を格納する。この状態は、パラメータ計算に用いられるため、パラメータ計算部 1091 に出力される。ベクトル符号帳 1094 は、コードベクトルを格納する。

次に、本発明に係るゲイン符号化方法のアルゴリズムについて説明する。

予め、量子化対象ベクトルの代表的サンプル（コードベクトル）が複数格納  
15 されたベクトル符号帳 1094 を作成しておく。各ベクトルは、ACゲイン、SCゲインの対数値に対応する値、及びSCの予測係数の調整係数の3つの要素からなる。

この調整係数は、以前のサブフレームの状態に応じて予測係数を調整する係数である。具体的には、この調整係数は、以前のサブフレームの状態が極端に  
20 大きな値又は極端に小さな値である場合に、その影響を小さくするように設定される。この調整係数は、多数のベクトルサンプルを用いた本発明者らが開発した学習アルゴリズムにより求めることが可能である。ここでは、この学習アルゴリズムについての説明は省略する。

例えば、有声音に多くの頻度で用いるコードベクトルは調整係数を大きく設  
25 定する。すなわち、同じ波形が並んでいる場合には、以前のサブフレームの状態の信頼性が高いので調整係数を大きくして、以前のサブフレームの予測係数をそのまま利用できるようにする。これにより、より効率的な予測を行うこと

ができる。

一方、語頭などに使用するあまり使用頻度の少ないコードベクトルは調整係数を小さくする。すなわち、前の波形と全然違う場合には、以前のサブフレーム

- 
- 5 係数を小さくして、以前のサブフレームの予測係数の影響を小さくする。これにより、次の予測の弊害を防いで良好な予測符号化を実現することができる。

このように、各コードベクトル（状態）に応じて予測係数を制御することにより、これまでの予測符号化の性能をさらに向上させることができる。

- また、予測係数格納部 1 0 9 5 には、予測符号化を行うための予測係数を格納しておく。この予測係数は MA (moving average) の予測係数で AC と SC の 2 種類を予測次数分格納する。これらの予測係数値は、一般に、予め多くのデータを用いた学習により求めておく。また、復号化ベクトル格納部 1 0 9 6 には、初期値として無音状態を示す値を格納しておく。

- 次に、符号化方法について詳細に説明する。まず、パラメータ計算部 1 0 9 1 に聴感重み付け入力音声 ( $X_i$ )、聴感重み付け LPC 合成済み適応音源 ( $A_i$ )、聴感重み付け LPC 合成済み確率的音源 ( $S_i$ ) を送り、さらに復号化ベクトル格納部 1 0 9 6 に格納された復号化ベクトル (AC、SC、調整係数)、予測係数格納部 1 0 9 5 に格納された予測係数 (AC、SC) を送る。これらを用いて符号化歪計算に必要なパラメータを計算する。

- 20 符号化歪計算部 1 0 9 2 における符号化歪計算は、下記式 2 にしたがって行う。

$$E_n = \sum_{i=0}^l (X_i - G_{an} \times A_i - G_{sn} \times S_i)^2$$

式 2

ここで、

- 25  $G_{an}$ ,  $G_{sn}$  : 復号化ゲイン

$E_n$  : n 番のゲインコードベクトルを用いたときの符号化歪み

$X_i$  : 聴感重み付け音声

$A_i$  : 聴感重み付けL P C合成済み適応音源

$S_i$  : 聴感重み付けL P C合成済み確率的音源

$n$  : コードベクトルの番号

5  $i$  : 音源ベクトルのインデクス

$I$  : サブフレーム長 (入力音声の符号化単位)

この場合、演算量を少なくするために、パラメータ計算部1091では、コードベクトルの番号に依存しない部分の計算を行う。計算しておくものは、上記予測ベクトルと3つの合成音 ( $X_i$ ,  $A_i$ ,  $S_i$ ) 間の相関、パワである。こ

10 の計算は、下記式3にしたがって行う。

$$D_{xx} = \sum_{i=0}^I X_i \times X_i$$

$$D_{xa} = \sum_{i=0}^I X_i \times A_i \times 2$$

$$D_{xs} = \sum_{i=0}^I X_i \times S_i \times 2$$

$$D_{aa} = \sum_{i=0}^I A_i \times A_i$$

$$D_{as} = \sum_{i=0}^I A_i \times S_i \times 2$$

$$D_{ss} = \sum_{i=0}^I S_i \times S_i$$

式3

$D_{xx}$ ,  $D_{xa}$ ,  $D_{xs}$ ,  $D_{aa}$ ,  $D_{as}$ ,  $D_{ss}$  : 合成音間の相関値、パワ

$X_i$  : 聴感重み付け音声

15  $A_i$  : 聴感重み付けL P C合成済み適応音源

$S_i$  : 聴感重み付けL P C合成済み確率的音源

$n$  : コードベクトルの番号

$i$  : 音源ベクトルのインデクス

$I$  : サブフレーム長 (入力音声の符号化単位)

また、パラメータ計算部1091では、復号化ベクトル格納部1096に格納された過去のコードベクトルと、予測係数格納部1095に格納された予測係数を用いて下記式4に示す3つの予測値を計算しておく。

$$\begin{aligned} P_{ra} &= \sum_{m=0}^M \alpha_m \times S_{am} \\ P_{rs} &= \sum_{m=0}^M \beta_m \times S_{cm} \times S_{sm} \\ P_{sc} &= \sum_{m=0}^M \beta_m \times S_{cm} \end{aligned}$$

5

式4

ここで、

$P_{ra}$  : 予測値 (ACゲイン)

$P_{rs}$  : 予測値 (SCゲイン)

$P_{sc}$  : 予測値 (予測係数)

10  $\alpha_m$  : 予測係数 (ACゲイン、固定値)

$\beta_m$  : 予測係数 (SCゲイン、固定値)

$S_{am}$  : 状態 (過去のコードベクトルの要素、ACゲイン)

$S_{sm}$  : 状態 (過去のコードベクトルの要素、SCゲイン)

$S_{cm}$  : 状態 (過去のコードベクトルの要素、SC予測係数調整係数)

15  $m$  : 予測インデクス

$M$  : 予測次数

上記式4から分かるように、 $P_{rs}$ 、 $P_{sc}$ については、従来と異なり調整係数が乗算されている。したがって、SCゲインの予測値及び予測係数については、調整係数により、以前のサブフレームにおける状態の値が極端に大きい小さい場合に、それを緩和する（影響を小さくする）ことができる。すなわち、状態に応じて適応的にSCゲインの予測値及び予測係数を変化させることが可能となる。

20

次に、符号化歪計算部1092において、パラメータ計算部1091で計算



した各パラメータ、予測係数格納部 1 0 9 5 に格納された予測係数、及びベクトル符号帳 1 0 9 4 に格納されたコードベクトルを用いて、下記式 5 にしたがって符号化歪を算出する。

$$E_n = D_{xx} + (G_{an})^2 \times D_{aa} + (G_{sn})^2 \times D_{ss} - G_{an} \times D_{xa} - G_{sn} \times D_{xs} + G_{an} \times G_{sn} \times D_{as}$$

$$G_{an} = P_{ra} + (1 - P_{ac}) \times C_{an}$$

$$G_{sn} = 10^{\{P_{rs} + (1 - P_{sc}) \times C_{sn}\}}$$

5

式 5

ここで、

$E_n$  :  $n$  番のゲインコードベクトルを用いたときの符号化歪み

$D_{xx}$ ,  $D_{xa}$ ,  $D_{xs}$ ,  $D_{aa}$ ,  $D_{as}$ ,  $D_{ss}$  : 合成音間の相関値、パワ

$G_{an}$ ,  $G_{sn}$  : 復号化ゲイン

10  $P_{ra}$  : 予測値 (ACゲイン)

$P_{rs}$  : 予測値 (SCゲイン)

$P_{ac}$  : 予測係数の和 (固定値)

$P_{sc}$  : 予測係数の和 (上記式 4 で算出)

$C_{an}$ ,  $C_{sn}$ ,  $C_{cn}$  : コードベクトル、 $C_{cn}$  は予測係数調整係数であるがこ

15 こでは使用しない

$n$  : コードベクトルの番号

なお、実際には  $D_{xx}$  はコードベクトルの番号  $n$  に依存しないので、その加算を省略することができる。

次いで、比較部 1 0 9 3 は、ベクトル符号帳 1 0 9 4 と符号化歪計算部 1 0  
20 9 2 を制御し、ベクトル符号帳 1 0 9 4 に格納された複数のコードベクトルの中で符号化歪計算部 1 0 9 2 にて算出された符号化歪みの最も小さくなるコードベクトルの番号を求め、これをゲインの符号とする。また、得られたゲインの符号を用いて復号化ベクトル格納部 1 0 9 6 の内容を更新する。更新は、

下記式 6 にしたがって行う。

$$S_{am} = S_{am-1}(M=M \sim 1), S_{a0} = C_a J$$

$$S_{sm} = S_{sm-1}(M=M \sim 1), S_{s0} = C_s J$$

$$S_{cm} = S_{cm-1}(M=M \sim 1), S_{c0} = C_c J$$

5

式 6

ここで、

$S_{am}$ ,  $S_{sm}$ ,  $S_{cm}$  : 状態ベクトル (AC、SC、予測係数調整係数)

$m$  : 予測インデクス

$M$  : 予測次数

10  $J$  : 比較部で求められた符号

式 4 から式 6 までで分かるように、本実施の形態では、復号化ベクトル格納部 1096 で状態ベクトル  $S_{cm}$  を格納しておいて、この予測係数調整係数を用いて予測係数を適応的に制御している。

図 5 は、本発明の実施の形態の音声復号化装置の構成を示すブロック図である。この音声復号化装置は、図 1 に示す音声復号化部 18 に含まれている。なお、図 5 に示す適応符号帳 202 は図 1 に示す RAM 22 に格納されており、図 5 に示す確率的符号帳 203 は図 1 に示す ROM 23 に格納されている。

図 5 に示す音声復号化装置において、パラメータ復号化部 201 は、伝送路から、符号化された音声信号を得ると共に、各音源符号帳 (適応符号帳 202、確率的符号帳 203) の音源サンプルの符号、LPC 符号、及びゲイン符号を得る。そして、LPC 符号から復号化された LPC 係数を得て、ゲイン符号から復号化されたゲインを得る。

そして、音源作成部 204 は、それぞれの音源サンプルに復号化されたゲインを乗じて加算することによって復号化された音源信号を得る。この際、得られた復号化された音源信号を、音源サンプルとして適応符号帳 204 へ格納し、同時に古い音源サンプルを破棄する。そして、LPC 合成部 205 では、復号化された音源信号に復号化された LPC 係数によるフィルタリングを行うこ

25

とによって、合成音を得る。

また、2つの音源符号帳は、図2に示す音声符号化装置に含まれるもの（図2の参照符号103、104）と同様のものであり、音源サンプルを取り出すためのサンプル番号（適応符号帳への符号と確率的符号帳への符号）は、いずれもパラメータ復号化部201から供給される。

このように、本実施の形態の音声符号化装置では、各コードベクトルに応じて予測係数を制御することが可能になり、音声の局所的特徴により適応したより効率的な予測や、非定常部における予測の弊害を防ぐことが可能になり、従来得られなかった格別の効果を得ることができる。

#### 10 （実施の形態2）

音声符号化装置において、上述したように、ゲイン演算部では、音源作成部から得られた適応符号帳、確率的符号帳の全ての音源について合成音と入力音声との間の比較を行う。このとき、演算量の都合上、通常は2つの音源（適応符号帳と確率的符号帳）はオープンループに探索される。以下、図2を参照して説明する。

このオープンループ探索においては、まず、音源作成部105は適応符号帳103からのみ音源候補を次々に選び、聴感重みLPC合成部106を機能させて合成音を得て、ゲイン演算部108へ送り、合成音と入力音声との間の比較を行って最適な適応符号帳103の符号を選択する。

20 次いで、上記適応符号帳103の符号を固定して、適応符号帳103からは同じ音源を選択し、確率的符号帳104からはゲイン演算部108の符号に対応した音源を次々に選択して聴感重みLPC合成部106へ伝送する。ゲイン演算部108で両合成音の和と入力音声との間の比較を行って確率的符号帳104の符号を決定する。

25 このアルゴリズムを用いた場合、全ての符号帳の符号をそれぞれに対して全て探索するよりは符号化性能は若干劣化するが、計算量は大幅に削減される。このため一般にはこのオープンループ探索が用いられる。

ここで、従来のオープンループの音源探索の中で代表的なアルゴリズムについて説明する。ここでは、1つの分析区間（フレーム）に対して2つのサブフレームで構成する場合の音源探索手順について説明する。

- まず、ゲイン演算部108の指示を受けて、音源作成部105は適応符号帳
- 5 103から音源を引出して聴感重みLPC合成部106へ送る。ゲイン演算部108において、合成された音源と第1サブフレームの入力音声との間の比較を繰り返して最適な符号を求める。ここで、適応符号帳の特徴を示す。適応符号帳は過去において合成に使用した音源である。そして、符号は、図6に示すようにタイムラグに対応している。
- 10 次に、適応符号帳103の符号が決まった後に、確率的符号帳の探索を行う。音源作成部105は適応符号帳103の探索で得られた符号の音源とゲイン演算部108で指定された確率的符号帳104の音源とを取り出して聴感重みLPC合成部106へ送る。そして、ゲイン演算部108において、聴感重み付け済みの合成音と聴感重み付け済みの入力音声との間の符号化歪みを計算し、最も適当な（二乗誤差が最小となるもの）確率的音源104の符号を決
- 15 める。1つの分析区間（サブフレームが2の場合）での音源符号探索の手順を以下に示す。
- 1) 第1サブフレームの適応符号帳の符号を決定
  - 2) 第1サブフレームの確率的符号帳の符号を決定

20 3) パラメータ符号化部109でゲインを符号化し、復号化ゲインで第1サブフレームの音源を作成し、適応符号帳103を更新する。

  - 4) 第2サブフレームの適応符号帳の符号を決定
  - 5) 第2サブフレームの確率的符号帳の符号を決定
  - 6) パラメータ符号化部109でゲインを符号化し、復号化ゲインで第2サ

25 ブフレームの音源を作成し、適応符号帳103を更新する。

上記アルゴリズムによって効率よく音源の符号化を行うことができる。しかしながら、最近では、さらなる低ビットレート化を目指し、音源のビット数を

節約する工夫が行われている。特に注目されているのは、適応符号帳のラグに大きな相関があることを利用して、第1サブフレームの符号はそのままで、第2サブフレームの探索範囲を第1サブフレームのラグの近くに狭めて（エントリ数を減らして）ビット数を少なくするというアルゴリズムである。

- 5      このアルゴリズムでは、分析区間（フレーム）の途中から音声に変化する場合や、2つのサブフレームの様子が大きく異なる場合には局所的劣化を引き起こすことが考えられる。

- 10      本実施の形態では、符号化の前に2つのサブフレーム両方についてピッチ分析を行って相関値を算出し、得られた相関値に基づいて2つのサブフレームのラグの探索範囲を決定する探索方法を実現する音声符号化装置を提供する。

- 15      具体的には、本実施の形態の音声符号化装置は、1つのフレームを複数のサブフレームに分解してそれぞれを符号化するCELP型符号化装置において、最初のサブフレームの適応符号帳探索の前に、フレームを構成する複数のサブフレームのピッチ分析を行って相関値を算出するピッチ分析部と、上記ピッチ分析部がフレームを構成する複数のサブフレームの相関値を算出すると共に、その相関値の大小から各サブフレームで最もピッチ周期らしい値（代表ピッチと呼ぶ）を求め、ピッチ分析部にて得られた相関値と代表ピッチとに基づいて複数のサブフレームのラグの探索範囲を決定する探索範囲設定部と、を備えることを特徴としている。そして、この音声符号化装置では、探索範囲設定部において、ピッチ分析部で得た複数のサブフレームの代表ピッチと相関値を利用して探索範囲の中心となる仮のピッチ（仮ピッチと呼ぶ）を求め、探索範囲設定部において、求めた仮ピッチの周りの指定の範囲にラグの探索区間を設定し、ラグの探索区間を設定するときに、仮ピッチの前後に探索範囲を設定する。また、その際に、ラグの短い部分の候補を少なくし、ラグのより長い範囲を広く
- 20      設定し、適応符号帳探索の際に上記探索範囲設定部で設定された範囲でラグの探索を行う。

以下、本実施の形態に係る音声符号化装置について添付図面を用いて詳細に

説明する。ここでは、1フレームは2サブフレームに分割されているものとする。3サブフレーム以上の場合でも同様の手順で符号化を行うことができる。

この音声符号化装置においては、いわゆるデルタラグ方式によるピッチ探索において、分割されたサブフレームについてすべてピッチを求め、ピッチ間で  
5    どの程度の相関があるかどうかを求めて、その相関結果に応じて探索範囲を決定する。

図7は、本発明の実施の形態2に係る音声符号化装置の構成を示すブロック図である。まず、LPC分析部302において、入力された音声データ（入力音声）301に対して自己相関分析とLPC分析を行うことによってLPC係  
10    数を得る。また、LPC分析部302において、得られたLPC係数の符号化を行ってLPC符号を得る。さらに、LPC分析部302において、得られたLPC符号を復号化して復号化LPC係数を得る。

次いで、ピッチ分析部310において、2サブフレーム分の入力音声のピッチ分析を行い、ピッチ候補とパラメータを求める。1サブフレームに対するアル  
15    ゴリズムを以下に示す。相関係数は、下記式7により、2つ求められる。なおこの時、 $C_{pp}$ は $P_{min}$ についてまず求め、あとの $P_{min+1}$ 、 $P_{min+2}$ については、フレーム端の値の足し引きで効率的に計算できる。

$$V_p = \sum_{i=0}^L X_i \times X_i - P \quad (P=P_{min} \sim P_{max})$$

$$C_{pp} = \sum_{i=0}^L X_i - P \times X_i - P \quad (P=P_{min} \sim P_{max})$$

20

式7

ここで、

$X_i$ ,  $X_{i-p}$  : 入力音声

$V_p$  : 自己相関関数

$C_{pp}$  : パワ成分

25     $i$  : 入力音声のサンプル番号

L : サブフレームの長さ

P : ピッチ

$P_{min}$ ,  $P_{max}$  : ピッチの探索を行う最小値と最大値

そして、上記式 7 で求めた自己相関関数とパワ成分はメモリに蓄えておき、

- 5 次の手順で代表ピッチ  $P_1$  を求める。これは  $V_p$  が正で  $V_p \times V_p / C_{pp}$  を最大にするピッチ  $P$  を求める処理となっている。ただし、割り算は一般的に計算量がかかるので、分子と分母を 2 つとも格納し、掛け算に直して効率化を図っている。

- 10 ここでは、入力音声と入力音声からピッチ分過去の適応音源との差分の二乗和が最も小さくなるようなピッチを探す。この処理は  $V_p \times V_p / C_{pp}$  を最大にするピッチ  $P$  を求める処理と等価となる。具体的な処理は以下ようになる。

1) 初期化 ( $P = P_{min}$ ,  $VV = C = 0$ ,  $P_1 = P_{min}$ )

2) もし ( $V_p \times V_p \times C < VV \times C_{pp}$ ) 又は ( $V_p < 0$ ) ならば 4) へ。それ以外なら 3) へ。

- 15 3)  $VV = V_p \times V_p$ ,  $C = C_{pp}$ ,  $P_1 = P$  として 4) へ

4)  $P = P + 1$  とする。この時  $P > P_{max}$  であれば終了、それ以外の場合には 2) へ。

- 20 上記作業を 2 サブフレームのそれぞれについて行い、代表ピッチ  $P_1$ ,  $P_2$  と自己相関係数  $V_{1p}$ ,  $V_{2p}$ , パワー成分  $C_{1pp}$ ,  $C_{2pp}$  ( $P_{min} < p < P_{max}$ ) を求める。

次に、探索範囲設定部 3 1 1 で適応符号帳のラグの探索範囲を設定する。まず、その探索範囲の軸となる仮ピッチを求める。仮ピッチはピッチ分析部 3 1 0 で求めた代表ピッチとパラメータを用いて行う。

- 25 仮ピッチ  $Q_1$ ,  $Q_2$  は以下の手順で求める。なお、以下の説明においてラグの範囲として定数  $Th$  (具体的には 6 程度が適当である) を用いる。また、相関値は上記式 7 で求めたものを用いる。

まず、 $P_1$  を固定した状態で  $P_1$  の付近 ( $\pm Th$ ) で相関の最も大きい仮ピッ

チ ( $Q_2$ ) を見つける。

1) 初期化 ( $p = P_1 - Th$ ,  $C_{max} = 0$ ,  $Q_1 = P_1$ ,  $Q_2 = P_1$ )

2) もし ( $V_{1p1} \times V_{1p1} / C_{1p1p1} + V_{2p} \times V_{2p} / C_{2pp} < C_{max}$ ) または ( $V_{2p} < 0$ ) ならば4) へ。それ以外なら3) へ。

5 3)  $C_{max} = V_{1p1} \times V_{1p1} / C_{1p1p1} + V_{2p} \times V_{2p} / C_{2pp}$ ,  $Q_2 = p$  として4) へ

4)  $p = p + 1$  として2) へ。ただし、この時  $p > P_1 + Th$  であれば5) へ。

10 このようにして2) ~ 4) の処理を  $P_1 - Th \sim P_1 + Th$  まで行って、相関の最も大きいもの  $C_{max}$  と仮ピッチ  $Q_2$  を求める。

次に、 $P_2$  を固定した状態で  $P_2$  の付近 ( $\pm Th$ ) で相関の最も大きい仮ピッチ ( $Q_1$ ) を求める。この場合、 $C_{max}$  は初期化しない。 $Q_2$  を求めた際の  $C_{max}$  を含めて相関が最大となる  $Q_1$  を求めることにより、第1, 第2サブフレーム間で最大の相関を持つ  $Q_1$ ,  $Q_2$  を求めることが可能となる。

15 5) 初期化 ( $p = P_2 - Th$ )

6) もし ( $V_{1p} \times V_{1p} / C_{1pp} + V_{2p2} \times V_{2p2} / C_{2p2p2} < C_{max}$ ) 又は ( $V_{1p} < 0$ ) ならば8) へ。それ以外は7) へ。

7)  $C_{max} = V_{1p} \times V_{1p} / C_{1pp} + V_{2p2} \times V_{2p2} / C_{2p2p2}$ ,  $Q_1 = p$ ,  $Q_2 = P_2$  として8) へ。

20 8)  $p = p + 1$  として6) へ。ただし、この時  $p > P_2 + Th$  であれば9) へ。

9) 終了。

25 このようにして6) ~ 8) の処理を  $P_2 - Th \sim P_2 + Th$  まで行って、相関の最も大きいもの  $C_{max}$  と仮ピッチ  $Q_1$ ,  $Q_2$  を求める。この時の  $Q_1$ ,  $Q_2$  が第1サブフレームと第2サブフレームの仮ピッチである。

上記アルゴリズムにより、2つのサブフレームの相関を同時に評価しながら大きさに比較的差のない (差の最大は  $Th$  である) 仮ピッチを2つ選択するこ



とができる。この仮ピッチを用いることにより、第2サブフレームの適応符号帳探索の際に、探索の範囲を狭く設定しても符号化性能を大きく劣化させることを防止できる。例えば、第2サブフレームから音質が急に変化した場合などで、第2サブフレームの相関が強い場合は、第2サブフレームの相関を反映した $Q_1$ を用いることで第2サブフレームの劣化を回避出来る。

さらに、探索範囲設定部311は、求めた仮ピッチ $Q_1$ を用いて適応符号帳の探索を行う範囲( $L_{ST} \sim L_{EN}$ )を下記式8のようにして設定する。

第1サブフレーム

$$L_{ST} = Q_1 - 5 \quad (\text{ただし、} L_{ST} < L_{min} \text{ の時 } L_{ST} = L_{min})$$

$$L_{EN} = L_{ST} + 20 \quad (\text{ただし、} L_{ST} > L_{max} \text{ の時 } L_{ST} = L_{max})$$

第2サブフレーム

$$L_{ST} = T_1 - 10 \quad (\text{ただし、} L_{ST} < L_{min} \text{ の時 } L_{ST} = L_{min})$$

$$L_{EN} = L_{ST} + 21 \quad (\text{ただし、} L_{ST} > L_{max} \text{ の時 } L_{ST} = L_{max})$$

式8

15      ここで、

$L_{ST}$  : 探索範囲の最小

$L_{EN}$  : 探索範囲の最大

$L_{min}$  : ラグの最小値 (例 : 20)

$L_{max}$  : ラグの最大値 (例 : 143)

20       $T_1$  : 第1フレームの適応符号帳ラグ

上記設定において、第1サブフレームは探索範囲を狭める必要はない。しかしながら、本発明者らは、入力音声のピッチに基づいた値の付近を探索区間とした方が性能が良いことを実験により確認しており、本実施の形態では26サンプルに狭めて探索するアルゴリズムを使用している。

25      また、第2サブフレームは第1サブフレームで求められたラグ $T_1$ を中心にその付近に探索範囲を設定している。したがって、合計32エントリで、第2サブフレームの適応符号帳のラグを5ビットで符号化できることになる。また、

本発明者らは、この時もラグの小さい候補を少なく、ラグの大きい候補を多く設定することにより、より良い性能が得られることを実験により確認している。ただし、これまでの説明でわかるように、本実施の形態においては、仮ピッチ  $Q_2$  は使用しない。

5       ここで、本実施の形態における効果について説明する。探索範囲設定部 3 1 1 によって得られた第 1 サブフレームの仮ピッチの近くには、第 2 サブフレームの仮ピッチも存在している（定数  $T_h$  で制限したため）。また、第 1 サブフレームにおいて探索範囲を絞って探索しているので、探索の結果得られるラグは第 1 サブフレームの仮ピッチから離れない。

10       したがって、第 2 サブフレームの探索の時には、第 2 サブフレームの仮ピッチから近い範囲を探索できることになり、第 1、第 2 サブフレームの両方において適当なラグが探索できることになる。

例として、第 1 サブフレームが無音で、第 2 サブフレームから音声が発生した場合を考える。従来法では、探索範囲を狭めることで第 2 サブフレームのピッチが探索区間に含まれなくなると、音質は大きく劣化してしまう。本実施の形態に係る方法においては、ピッチ分析部の仮ピッチの分析において、代表ピッチ  $P_2$  の相関は強く出る。したがって、第 1 サブフレームの仮ピッチは  $P_2$  付近の値になる。このため、デルタラグによる探索の際に、音声が発生した部分に近い部分を仮ピッチとすることができる。すなわち、第 2 サブフレームの適応符号帳の探索の時には、 $P_2$  付近の値を探索できることになり、途中で音声の立ち上がり生じて劣化なくデルタラグにより第 2 サブフレームの適応符号帳探索を行うことができる。

次に、音源作成部 3 0 5 において、適応符号帳 3 0 3 に格納された音源サンプル（適応コードベクトル又は適応音源）と確率的符号帳 3 0 4 に格納された音源サンプル（確率的コードベクトル又は確率的音源）を取り出し、それぞれを聴感重み L P C 合成部 3 0 6 へ送る。さらに、聴感重み L P C 合成部 3 0 6 において、音源作成部 3 0 5 で得られた 2 つの音源に対して、L P C 分析部 3

02で得られた復号化LPC係数によってフィルタリングを行って2つの合成音を得る。

さらに、ゲイン演算部308においては、聴感重みLPC合成部306で得られた2つの合成音と入力音声との関係を分析し、2つの合成音の最適値（最適ゲイン）を求める。また、ゲイン演算部308においては、その最適ゲインによってパワ調整したそれぞれの合成音を加算して総合合成音を得る。そして、ゲイン演算部308は、その総合合成音と入力音声の符号化歪みの計算を行う。また、ゲイン演算部308においては、適応符号帳303と確率的符号帳304の全ての音源サンプルに対して音源作成部305、聴感重みLPC合成部306を機能させることによって得られる多くの合成音と入力音声との間の符号化歪みを行い、その結果得られる符号化歪みの中で最も小さいときの音源サンプルのインデクスを求める。

次に、得られた音源サンプルのインデクス、そのインデクスに対応する2つの音源、及び入力音声をパラメータ符号化部309へ送る。パラメータ符号化部309では、ゲインの符号化を行うことによってゲイン符号を得て、LPC符号、音源サンプルのインデクスと共に伝送路へ送る。

また、パラメータ符号化部309は、ゲイン符号と音源サンプルのインデクスに対応する2つの音源から実際の音源信号を作成し、それを適応符号帳303に格納すると同時に古い音源サンプルを破棄する。

なお、聴感重みLPC合成部306においては、LPC係数や高域強調フィルタや長期予測係数（入力音声の長期予測分析を行うことによって得られる）を用いた聴感重み付けフィルタを用いる。

上記ゲイン演算部308は、音源作成部305から得られた適応符号帳303、確率的符号帳304の全ての音源について入力音声との間の比較を行うが、計算量削減のため、2つの音源（適応符号帳303と確率的符号帳304）については上述したようにしてオープンループにより探索する。

このように、本実施の形態におけるピッチ探索方法により、最初のサブフレ

ームの適応符号帳探索の前に、フレームを構成する複数のサブフレームのピッチ分析を行って相関値を算出することにより、フレーム内の全サブフレームの相関値を同時に把握することができる。

- そして、各サブフレームの相関値を算出すると共に、その相関値の大小から
- 5 各サブフレームで最もピッチ周期らしい値（代表ピッチと呼ぶ）を求め、ピッチ分析で得られた相関値と代表ピッチに基づいて複数のサブフレームのラグの探索範囲を設定する。この探索範囲の設定においては、ピッチ分析で得た複数のサブフレームの代表ピッチと相関値を利用して探索範囲の中心となる差の少ない適当な仮のピッチ（仮ピッチと呼ぶ）を求める。
  - 10 さらに、上記探索範囲の設定で求めた仮ピッチの前後の指定の範囲にラグの探索区間を限定するので、適応符号帳の効率の良い探索を可能にする。その際、ラグの短い部分の候補を少なくし、ラグのより長い範囲を広く設定するので、良好な性能が得られる適当な探索範囲を設定することができる。また、適応符号帳探索の際に上記探索範囲の設定で設定された範囲でラグの探索を行うの
  - 15 で、良好な復号化音を得ることができる符号化が可能になる。

- このように、本実施の形態によれば、探索範囲設定部 311 によって得られた第 1 サブフレームの仮ピッチの近くには第 2 サブフレームの仮ピッチも存在しており、第 1 サブフレームにおいて探索範囲を絞っているので、探索の結果得られるラグは仮ピッチから離れて行かない。したがって、第 2 サブフレームの探索の時には第 2 サブフレームの仮ピッチ付近を探索できることになり、フレームの後半から音声が始まる場合などの非定常なフレームでも、第 1、第 2 サブフレームにおいて適当なラグ探索が可能になり、従来得られなかった格別の効果を得ることができる。

（実施の形態 3）

- 25 初期の CELP 方式では、ランダム数列が確率的音源ベクトルとして複数種類エンタリーされた確率的符号帳、すなわち複数種類のランダム数列をメモリに直接記録した確率的符号帳が使われていた。一方、近年の低ビットレート C

ELP 符号化・復号化装置においては、振幅が+1か-1の非零要素（非零要素以外の要素の振幅は零）を少数個含んだ確率的音源ベクトルを生成する代数的符号帳を確率的符号帳部に備えるものが多く開発されている。

- 5     なお、代数的符号帳は、「Fast CELP Coding based on Algebraic codes」, J.Adoul et al, Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing, 1987, pp. 1957-1960や「Comparison of Some Algebraic Structure for CELP Coding of Speech」, J.Adoul et al, Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing, 1987, pp. 1953-1956などに開示されている。

- 10    上記文献に開示されている代数的符号帳は、（1）ビットレートが8 kb/s程度のCELP方式に適用した場合、品質の高い合成音を生成できる、（2）少ない演算量で確率的音源符号帳を探索できる、（3）確率的音源ベクトルを、直接格納しておくデータROM容量が不要になる、といった優れた特徴を有する符号帳である。

- 15    そして、代数符号帳を確率的符号帳として用いることを特徴とするCS-ACELP（ビットレート8 kb/s）やACELP（ビットレート5.3 kb/s）が、G. 729、g 723.1として、それぞれITU-Tから1996年に勧告化されている。なお、CS-ACELPに関しては、「Design and Description of CS-ACELP: A Toll Quality 8 kb/s Speech Coder」, Redwan Salami et al, IEEE trans. SPEECH AND AUDIO PROCESSING, 20    vol. 6, no. 2, March 1998などに、その詳細技術が開示されている。

- 代数的符号帳は、上記のように優れた特徴を有する符号帳である。しかしながら、代数的符号帳をCELP符号化・復号化装置の確率的符号帳に適用した場合、確率的音源ターゲットは、比零要素を少数個だけ含んだ確率的音源ベクトルで常に符号化（ベクトル量子化）されることになるので、確率的音源ターゲットの忠実な符号表現は不可能であるという課題も生じている。そして、処理フレームが、無声子音区間や背景雑音区間などに相当する場合に、この課題
- 25

は特に顕著になる。

無声子音区間や背景雑音区間では、確率的音源ターゲットが複雑な形状になることが多いためである。またさらには、ビットレートが8 kb/s 程度よりさらに低いCELP符号化・復号化装置に代数的符号帳を適用した場合には、確率的音源ベクトル中の比零要素数を少なくすることになるため、確率的音源ターゲットがパルスの形状になりやすい有声区間でさえも、上記課題が問題になる場合がある。

代数的符号帳の有する上記課題を解決する一方法として、代数的符号帳より出力される少数個の非零要素（非零要素以外の要素はゼロの値を持つ）を含むベクトルと、拡散パターンと呼ばれる固定波形とを重畳して得られるベクトルを、合成フィルタの駆動音源とするパルス拡散符号帳を用いる方法が開示されている。パルス拡散符号帳は、特開平10-232696号公報、「パルス拡散構造音源を併用するACELP符号化」安永他，電子情報通信学会平成9年度春季全国大会発表予稿集，D-14-11，p. 253，1997-03、「パルス拡散音源を用いた低レート音声符号化」安永他，日本音響学会平成10年秋期研究発表会講演論文集，pp. 281-282，1998-10などに開示されている。

そこで次に、上記文献で開示されたパルス拡散符号帳の概要を、図8及び図9を用いて説明する。なお、図9は、図8のパルス拡散符号帳のさらに詳細な一例を示すものである。

図8及び図9のパルス拡散符号帳において、代数的符号帳4011は、少数個の非零要素（振幅は+1又は-1）からなるパルスベクトルを生成する符号帳である。上記文献に記載されているCELP符号化装置・復号化装置では、代数的符号帳4011の出力であるパルスベクトル（少数個の非零要素によって構成される）がそのまま、確率的音源ベクトルとして用いられている。

拡散パターン格納部4012には、拡散パターンと呼ばれる固定波形が、各チャネルあたり1種類以上ずつ格納されている。なお、各チャネルごとに格納された前記拡散パターンは、チャネル毎で異なる形状の拡散パターンが格納される場合、

各チャネルに同一形状（共通の）の拡散パターンが格納される場合の双方が考えられる。各チャネル用に格納される拡散パターンが共通の場合は、各チャネル用に格納される拡散パターンが格納される場合を単純化したものに相当するので、本明細書の以下の説明では、チャネル毎に格納される拡散パタンの形状がそれぞれ異なる場合について説明を進めることとする。

パルス拡散符号帳 401 は、代数的符号帳 4011 からの出力ベクトルをそのまま確率的音源ベクトルとして出力するのではなく、代数的符号帳 4011 から出力されるベクトルと、拡散パターン格納部 4012 から読み出される拡散パターンとを、パルス拡散部 4013 でチャネルごとに重畳し、重畳演算によって得られるベクトルを加算して得られるベクトルを確率的音源ベクトルとして利用する。

なお、上記文献において開示されている CELP 符号化・復号化装置は、符号化装置と復号化装置で同一構成（代数的符号帳部のチャネル数、拡散パターン格納部に登録されている拡散パタンの種類数および形状などが、符号化装置側と復号化装置側で共通）のパルス拡散符号帳を用いることを特徴としている。そして、拡散パターン格納部 4012 に登録しておく拡散パタンの形状、種類数、複数種類以上登録している場合にはそれらの選択方法を効率的に設定することによって、合成音声の品質を向上を図っている。

なお、パルス拡散符号帳に関するここでの説明は、少数個の非零要素からなるパルスベクトルを生成する符号帳として、非零要素の振幅を +1 もしくは -1 に限定した代数的符号帳を用いた場合についての説明であるが、当該パルスベクトルを生成する符号帳としては、非零要素の振幅を限定しないマルチパルス符号帳や、レギュラーパルス符号帳を用いることも可能であり、その場合にも、パルスベクトルを拡散パターンと重畳したものを確率的音源ベクトルとして利用することで合成音声の品質向上を実現できる。

これまでに、多くの確率的音源ターゲットの形状を統計学習し、確率的音源ターゲット中に統計的に高い頻度で含まれる形状の拡散パターン、無声子音区間

や雑音区間を効率的に表現するための乱数的な形状の拡散パターン、有声定常区間を効率的に表現するためのパルス的な形状の拡散パターン、代数的符号帳から出力されるパルスベクトルのエネルギー（非零要素の位置にエネルギーが集中

- している）を周囲に分散させるような作用を与える形状の拡散パターン、適当に
- 5 用意したいくつかの拡散パターン候補について、音声信号を、符号化、復号化、合成音声の視聴評価を繰り返し、品質の高い合成音声を出力しうるよう選択した拡散パターン、又は音声学的な知見をもとに作成した拡散パターンなどを、代数的符号帳から出力される音源ベクトル中の非零要素（チャンネル）あたり1種類以上ずつ登録しておき、登録しておいた拡散パターンと、代数的符号帳によって
- 10 生成されるベクトル（少数個の非零要素によって構成される）とをチャンネルごとに重畳し、各チャンネルの重畳結果を加算したものを確率的音源ベクトルとして用いることにより、合成音声を有効に品質向上させることができることが開示されてきた。

- また、特に、拡散パターン格納部4012が、チャンネルあたり複数種類（2種類以上）の拡散パターンを登録している場合については、それら複数の拡散パ
- 15 ターの選択方法として、登録された拡散パターの全組合わせについて実際に符号化・復号化を行い、その結果生じる符号化歪みが最小になるような拡散パターンをクローズド選択する方法や、確率的符号帳探索を行う時点で既に明らかになっている音声的情報（ここでいう音声的情報とは、例えば、ゲイン符号の動的
- 20 変動情報もしくはゲイン値の（予め設定したしきい値との）大小関係情報などを利用して判定した有声性の強弱情報、あるいは、線形予測符号の動的変動を利用して判定した有声性の強弱情報などのことである）を利用して、拡散パターンをオープン選択する方法などが開示されている。

- なお、以降の説明では、説明を簡単にするため、図9のパルス拡散符号帳内の拡散パターン格納部4013が、チャンネルあたり1種類だけの拡散パターンを登録していることを特徴とする図10のパルス拡散符号帳に限定して説明する。
- 25

ここでは次に、代数的符号帳をCELP符号化装置に適用した場合の確率的



符号帳探索処理と比較して、パルス拡散符号帳をCELP符号化装置に適用した場合の確率的符号帳探索処理を説明する。まず、代数的符号帳を確率的符号帳部に用いた場合の符号帳探索処理を説明する。

- 5 代数的符号帳によって出力されるベクトル内の非零要素数をN（代数的符号帳のチャンネル数をN）、チャンネルごとに出力する振幅が+1か-1の非零要素を1本だけ含むベクトル（非零要素以外の要素の振幅はゼロ）を $d_i$ （ $i$ はチャンネル番号： $0 \leq i \leq N-1$ ）、サブフレーム長をLとした時、代数的符号帳によって出力されるエン트리番号 $k$ の確率的音源ベクトル $c_k$ は、下記式9となる。

10

$$C_k = \sum_{i=0}^{N-1} d_i$$

$C_k$ : 代数的符号帳によるエン트리番号 $K$ の確率的音源ベクトル

$d_i$ : 非零要素ベクトル（ $d_i = \pm \delta(n - p_i)$ ）ただし、 $p_i$ : 非零要素位置

$N$ : 代数的符号帳のチャンネル数 (= 確率的音源ベクトル中の非零要素数)

15

式9

そして、式9を式10に代入することで、下記式11が得られる。

$$D_k = \frac{(v^t H c_k)^2}{\|H c_k\|^2}$$

$v^t$  :  $v$  (確率的音源ターゲット) の転置ベクトル

- 20  $H^t$  :  $H$  (合成フィルタのインパルス応答行列) の転置行列

$c_k$ : エン트리番号 $k$ 番目の確率的音源ベクトル

式10

$$D_k = \frac{\left( \mathbf{v}^t \mathbf{H} \left( \sum_{i=0}^{N-1} \mathbf{d}_i \right) \right)^2}{\left\| \mathbf{H} \left( \sum_{i=0}^{N-1} \mathbf{d}_i \right) \right\|^2}$$

$\mathbf{v}$ : 確率的音源ターゲットベクトル

$\mathbf{H}$ : 合成フィルタのインパルス応答畳み込み行列

$\mathbf{d}_i$ : 非零要素ベクトル ( $\mathbf{d}_i = \pm \delta(n - p_i)$  ただし、 $p_i$ : 非零要素位置)

5  $N$ : 代数的符号帳のチャネル数 (= 確率的音源ベクトル中の非零要素数)

$$\mathbf{x}^t = \mathbf{v}^t \mathbf{H}$$

$$\mathbf{M} = \mathbf{H}^t \mathbf{H}$$

式 1 1

この式 1 0 を整理して得られる下記式 1 2 を最大化するようなエントリ番号  $k$  を特定する処理が確率的符号帳探索処理となる。

$$D_k = \frac{\left( \sum_{i=0}^{N-1} \mathbf{x}^t \mathbf{d}_i \right)^2}{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \mathbf{d}_i^t \mathbf{M} \mathbf{d}_j}$$

式 1 2

ただし、式 1 2 において、 $\mathbf{x}^t = \mathbf{v}^t \mathbf{H}$ 、 $\mathbf{M} = \mathbf{H}^t \mathbf{H}$  ( $\mathbf{v}$  は確率的音源ターゲット) である。ここで各エントリ番号  $k$  について式 1 2 の値を計算する場合、その前  
15 処理段階で  $\mathbf{x}^t = \mathbf{v}^t \mathbf{H}$  及び  $\mathbf{M} = \mathbf{H}^t \mathbf{H}$  を計算し、計算結果をメモリに展開 (記憶) させておく。この前処理を導入することで、確率的音源ベクトルとしてエントリしている各候補ごとに式 1 2 を計算する際の演算量を大幅に削減でき、この結果として、確率的符号帳探索に要するトータルの演算量を少なくおさえられることが、上記文献などに開示されており、一般に知られている。

20 次に、パルス拡散符号帳を確率的符号帳に用いた場合の確率的符号帳探索処理を説明する。

- パルス拡散符号帳の構成一部位である代数的符号帳によって出力される非  
 零要素数を $N$ （代数的符号帳のチャンネル数を $N$ ）、チャンネルごとに出力する振  
 幅が $+1$ か $-1$ の非零要素を1本だけ含むベクトル（非零要素以外の要素の振  
 幅はゼロ）を $d_i$ （ $i$ はチャンネル番号： $0 \leq i \leq N-1$ ）、拡散パターン格納部  
 5 が格納しているチャンネル番号 $i$ 用の拡散パターンを $w_i$ 、サブフレーム長を $L$ と  
 した時、パルス拡散符号帳によって出力されるエン트리番号 $k$ の確率的音源  
 ベクトル $c_k$ は、下記式13となる。

$$C_k = \sum_{i=0}^{N-1} W_i d_i$$

- 10  $C_k$ : パルス拡散符号帳によるエン트리番号 $K$ の確率的音源ベクトル

$W_i$ : 拡散パターン( $w_i$ )重畳行列

$d_i$ : 代数的符号帳部が出力する非零要素ベクトル

( $d_i = \pm \delta(n - p_i)$  ただし、 $p_i$ : 非零要素位置)

$N$ : 代数的符号帳部のチャンネル数

15

式13

したがってこの場合、式13を式10に代入することで、下記式14が得ら  
 れる。

$$D_k = \frac{\left( v^T H \left( \sum_{i=0}^{N-1} W_i d_i \right) \right)^2}{\left\| H \left( \sum_{i=0}^{N-1} W_i d_i \right) \right\|^2}$$

$v$ : 確率的音源ターゲットベクトル

- 20  $H$ : 合成フィルタのインパルス応答畳み込み行列

$W_i$ : 拡散パターン( $w_i$ )重畳行列

$d_i$ : 代表的符号帳部が出力する非零要素ベクトル

( $d_i = \pm \delta(n - p_i)$  ただし、 $p_i$ : 非零要素位置)

$N$ : 代数的符号帳のチャンネル数 (= 確率的音源ベクトル中の非零要素数)

$$H_i = HW_i$$

$$x_i^t = v^t H_i$$

$$R = H_i H_j$$

式 1 4

- 5 この式 1 4 を整理して得られる下記式 1 5 を最大化する確率的音源ベクトルのエントリ番号  $k$  を特定する処理が、パルス拡散符号帳を用いた場合の確率的符号帳探索処理となる。

$$D_k = \frac{\left( \sum_{i=0}^{N-1} x_i^t d_i \right)^2}{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} d_i^t R d_j}$$

式 1 5

- 10 ただし、式 1 5 において、 $x^t = v^t H_i$  (ただし、 $H_i = HW_i$  :  $W_i$  は拡散パターン重畳行列)、である。各エントリ番号  $k$  について式 1 5 の値計算する場合、その前処理として  $H_i = HW_i$  及び  $x^t = v^t H_i$  及び  $R = H_i^t H_j$  を計算しメモリに記録しておくことが可能である。すると、確率的音源ベクトルとしてエントリしている各候補ごとに式 1 5 を計算する際の演算量が、代数的符号帳を用いた場合に
- 15 式 1 2 を計算する際の演算量と同じになり (式 1 2 と式 1 5 が同形であることから明らか)、パルス拡散符号帳を用いた場合も、少ない演算量で確率的符号帳探索を行うことができる。

- 上記技術においては、パルス拡散符号帳を CELP 符号化装置・復号化装置の確率的符号帳部に用いることの効果、及びパルス拡散符号帳を確率的符号帳
- 20 部に用いた場合に、代数的符号帳を確率的符号帳部に用いた場合と同様の方法で確率的符号帳探索を行えることを示した。代数的符号帳を確率的符号帳部に用いた場合の確率的符号帳探索に要する演算量と、パルス拡散符号帳を確率的符号帳部に用いた場合の確率的符号帳探索に要する演算量の違いは、式 1 2 と式 1 5 それぞれの前処理段階に要する演算量の違い、すなわち、前処理

( $\mathbf{x}^t = \mathbf{v}^t \mathbf{H} \mathbf{i}$ ,  $\mathbf{M} = \mathbf{H}^t \mathbf{H}$ ) と前処理 ( $\mathbf{H} \mathbf{i} = \mathbf{H} \mathbf{W} \mathbf{i}$ ,  $\mathbf{x}^t = \mathbf{v}^t \mathbf{H} \mathbf{i}$ ,  $\mathbf{R} = \mathbf{H} \mathbf{i}^t \mathbf{H} \mathbf{j}$ ) に要する演算量の違いである。

一般に、C E L P 符号化装置・復号化装置では、そのビットレートが低くなるほど確率的符号帳部に割り当て可能なビット数も減少する傾向にある。そしてこの傾向は、代数的符号帳やパルス拡散符号帳を確率的符号帳部に用いる場合、確率的音源ベクトルを構成する際の子要素数の減少につながっていく。したがって、C E L P 符号化装置・復号化装置のビットレートが低くなるほど、代数的符号帳を用いた場合とパルス拡散符号帳を用いた場合の演算量の差は少なくなる。しかしビットレートが比較的高い場合や、ビットレートが低くても演算量を極力少なく押さえる必要がある場合には、パルス拡散符号帳を用いることによって生じる前処理段階の演算量の増加が無視できなくなることがある。

本実施の形態では、パルス拡散符号帳を確率的符号帳部に用いた C E L P 方式の音声符号化装置と音声復号化装置、及び音声符号化復号化システムにおいて、代数的符号帳を確率的符号帳部に用いる場合と比べて増加する、符号探索処理における前処理部分の演算量増加分を少なく抑えながら、復号化側では高品質な合成音声を得ることについて説明する。

具体的には、本実施の形態に係る技術は、パルス拡散符号帳を C E L P 符号化装置・復号化装置の確率的符号帳部に用いる場合に生じることがある上記課題を解決するためのものであり、符号化装置側と復号化装置側で異なる拡散パターンを用いることを特徴である。すなわち、本実施の形態においては、音声復号化装置側の拡散パターン格納部には、上述した拡散パターンを登録し、それを用いることで、代数的符号帳を用いる場合より品質の高い合成音声を生成する。一方、音声符号化装置側では、復号化装置側の拡散パターン格納部に登録する拡散パターンを簡素化した拡散パターン（例えば、一定間隔で間引いた拡散パターンや、ある長さで打ち切った拡散パターン）を登録し、それを用いて確率的符号帳探索を行うようにする。

これにより、パルス拡散符号帳を確率的符号帳部に用いる場合に、符号化側では、代数的符号帳を確率的符号帳部に用いる場合と比べて増加する、前処理段階の符号探索時の演算量を少なく抑えることができ、復号化側では、高品質の合成音声を得ることができる。

- 5      符号化装置側と復号化装置側で異なる拡散パターンを用いることとは、予め用意された（復号化装置用の）拡散ベクトルを、その特性を残しつつ変形することにより、エンコーダ用の拡散ベクトルを獲得することである。

- 10      ここで、復号化装置用の拡散ベクトルを予め用意する方法としては、本発明者らが以前に出願した特許（特開平10-63300号公報）に開示された方法、すなわち音源探索用ターゲットベクトルの統計的傾向を学習することによって用意する方法、音源ターゲットを実際に符号化し、その時生じる符号化歪みの総和をより小さくする方向に徐々に変形させる操作を反復することで用意する方法、及び合成音声を高品質化すべく音声学的な知見に基づいて設計する方法などや、パルス音源の高域位相成分をランダム化させることを目的に  
15      設計する方法などが考えられる。これらの内容はすべてここに含めておく。

- 20      このようにして得られた拡散ベクトルは、いずれも拡散ベクトルの先頭サンプルに近いサンプル（前方のサンプル）の振幅が、後方のサンプルの振幅より、比較的大きめになるという特徴がある。中でも、先頭のサンプルの振幅が、拡散ベクトル内の全サンプル中で最大となることが多い（ほとんどの場合そのようになる）。

復号化装置用の拡散ベクトルを、その特性を残しつつ変形することでエンコーダ用の拡散ベクトルを獲得する具体的方法としては、以下の方法が挙げられる。

- 25      1）復号化装置用の拡散ベクトルのサンプル値を、適当な間隔ごとにゼロに置き換えることで、エンコーダ用の拡散ベクトルを獲得する。

2）ある長さの復号化装置用の拡散ベクトルを、適当な長さで打ち切ることによって、エンコーダ用の拡散ベクトルを獲得する。

3) 振幅のしきい値を予め設定し、復号化装置用の拡散ベクトルに対して設定したしきい値より振幅の小さいサンプルをゼロに置き換えることで、エンコーダ用の拡散ベクトルを獲得する。

4) ある長さの復号化装置用の拡散ベクトルを、先頭サンプルを含む適当な  
5 間隔ごとのサンプル値を保存し、それ以外のサンプルの値をゼロに置きかえることで、符号化装置用の拡散ベクトルを獲得する。

ここで例えば上記1)の方法のように、拡散ベクトルの前方からの数サンプルを用いた場合でも、拡散ベクトルの概形（大まかな特性）を保存したまま、符号化装置用の拡散ベクトルを新たに獲得することが可能となっている。

10 また例えば、上記2)の方法のように、適当な間隔ごとにサンプル値をゼロに置き換えてももとの拡散ベクトルの概形（大まかな特性）を保存したまま、符号化装置用の拡散ベクトルを新たに獲得することが可能となる。特に、上記4)の方法の場合は、振幅が最大であることの多い先頭サンプルの振幅をそのまま必ず保存するという限定を付けているので、もとの拡散ベクトルの概形を  
15 より確実に保存しておくことが可能である。

また、3)の方法のように、特定値以上の振幅を有するサンプルをそのまま保存し、前記特定値以下の振幅を有するサンプルの振幅をゼロに置き換えても拡散ベクトルの概形（大まかな特性）を保存したまま、符号化装置用の拡散ベクトルを獲得することが可能となる。

20 以下、本実施の形態に係る音声符号化装置及び音声復号化装置について、添付図面を参照して詳細に説明する。なお、添付図面に記載のCELP音声符号化装置（図11）、および、CELP音声復号化装置（図12）は、従来のCELP音声装置およびCELP音声復号化装置における確率的符号帳部分に、上記のパルス拡散符号帳を用いている点に特徴を有している。従って、以降の  
25 説明において、確率的符号帳、確率的音源ベクトル、確率的音源ゲインと記載された部分は、それぞれ、パルス拡散符号帳、パルス拡散音源ベクトル、パルス拡散音源ゲインと読み替えることが可能である。なお、CELP音声符号化

装置およびCELP音声復号化装置における確率的符号帳は、雑音符号帳、あるいは、複数種類の固定波形を格納する作用を有することから固定符号帳と呼ばれることもある。

- 5        図11のCELP音声符号化装置では、まず始めに、線形予測分析部501が、入力音声線を形予測分析して線形予測係数を算出し、算出した線形予測係数を線形予測係数符号化部502へ出力する。次に、線形予測係数符号化部502が、線形予測係数を符号化（ベクトル量子化）し、ベクトル量子化によって得られる量子化インデックス（以下、線形予測符号と呼ぶ）を符号出力部513及び線形予測符号復号化部503へ出力する。

次いで、線形予測符号復号化部503が、線形予測係数符号化部502で得た線形予測符号を復号化（逆量子化）して合成フィルタ504へ出力する。合成フィルタ504は、線形予測符号復号化部503で復号化して得られた復号化線形予測符号を係数に持つ全極型モデルの合成フィルタを構成する。

- 15        そして、適応符号帳506から選出される適応音源ベクトルに適応音源ゲイン509を乗じて得られるベクトルと、パルス拡散符号帳507から選出した確率的音源ベクトルに確率的音源ゲイン510を乗じて得られるベクトルとをベクトル加算部511で加算して駆動音源ベクトルを生成する。そして、歪み計算部505が、当該駆動音源ベクトルで合成フィルタ504を駆動したときの出力ベクトルと、入力音声との歪みを下記式16により計算し、歪みERを符号特定部512へ出力する。

$$ER = \|u - (g_a H p + g_c H c)\|^2$$

u: 入力音声 (ベクトル)

H: 合成フィルタのインパルス応答行列

- 25        p: 適応音源ベクトル

c: 確率的音源ベクトル



$g_a$ : 適応音源ゲイン

$g_c$ : 確率的音源ゲイン

式 1 6

ただし、式 1 6 において、 $u$  は処理フレーム内の入力音声ベクトル、 $H$  は合成フィルタのインパルス応答行列、 $g_a$  は適応音源ゲイン、 $g_c$  は確率的音源ゲイン、 $p$  は適応音源ベクトル、 $c$  は確率的音源ベクトルを示す。

ここで、適応符号帳 5 0 6 は、過去数フレーム分の駆動音源ベクトルを格納したバッファ（動的メモリ）であり、上記適応符号帳 5 0 6 から選出される適応音源ベクトルは、入力音声を合成フィルタの逆フィルタに通して得られる線形予測残差ベクトル中の周期成分を表現するために使われる。

一方、パルス拡散符号帳 5 0 7 から選出される音源ベクトルは、線形予測残差ベクトルに現処理フレームで新たに加わった非周期成分（線形予測残差ベクトルから周期性（適応音源ベクトル成分）を除去した成分）を表現するために使われる。

そして、適応音源ベクトルゲイン乗算部 5 0 9 及び確率的音源ベクトルゲイン乗算部 5 1 0 は、適応符号帳 5 0 6 から選出される適応音源ベクトル及びパルス拡散符号帳 5 0 7 から選出される確率的音源ベクトルに対して、ゲイン符号帳 5 0 8 から読みだした適応音源ゲイン及び確率的音源ゲインを乗じる機能を有している。なお、ゲイン符号帳 5 0 8 とは、適応音源ベクトルに乘じる適応音源ゲインと、確率的音源ベクトルに乘じる確率的音源ゲインとのセットを複数種類格納した静的メモリである。

符号特定部 5 1 2 は、歪み計算部 5 0 5 で計算した式 1 6 の歪み  $ER$  を最小化する上記 3 つの符号帳（適応符号帳、パルス拡散符号帳、ゲイン符号帳）のインデクスの最適組み合わせを選択する。そして、歪み特定部 5 1 2 は、上記歪みが最小になるときに選択していた各符号帳のインデクスを、それぞれ適応音源符号、確率的音源符号、ゲイン符号として符号出力部 5 1 3 へ出力する。

そして最後に、符号出力部 5 1 3 は、線形予測係数符号化部 5 0 2 で得られ

た線形予測符号と、符号特定部 5 1 2 で特定された適応音源符号、確率的音源符号及びゲイン符号を、全てまとめて現処理フレーム内の入力音声を表現する符号（ビット情報）とし、復号化装置側へ出力する。

- 5      なお、符号特定部 5 1 2 で行う適応音源符号、確率的音源符号、ゲイン符号の特定は、一定時間間隔のフレームを、サブフレームと呼ぶさらに短い時間間隔に分割した上で行われることがある。ただし、本明細書では、フレームとサブフレームと特に区別しないで（フレームという呼び方に統一した上で）、以下の説明を行う。

次に、CELP 音声復号化装置の概要を、図 1 2 を用いて説明する。

- 10      図 1 2 の CELP 復号化装置では、まず、符号入力部 6 0 1 が、CELP 音声符号化装置（図 1 1）で特定した符号（フレーム区間内の音声信号を符号表現するためのビット情報）を受け、受けた符号を線形予測符号、適応音源符号、確率的音源符号、及びゲイン符号の 4 種類の符号に分解する。そして、線形予測符号を線形予測係数復号化部 6 0 2 へ、適応音源符号を適応符号帳 6 0 3 へ、  
15      確率的音源符号をパルス拡散符号帳 6 0 4 へ、ゲイン符号をゲイン符号帳 6 0 5 へ出力する。

次に、線形予測係数復号化部 6 0 2 は、符号入力部 6 0 1 から入力される線形予測符号を復号化して復号化線形予測符号を得て、この復号化線形予測符号を合成フィルタ 6 0 9 へ出力する。

- 20      合成フィルタ 6 0 9 は、線形予測係数復号化部 6 0 2 で得た復号化線形予測符号を係数にもつ全極型モデルの合成フィルタを構成する。また、適応符号帳 6 0 3 は、符号入力部 6 0 1 から入力された適応音源符号に対応する適応音源ベクトルを出力する。また、パルス拡散符号帳 6 0 4 は、符号入力部 6 0 1 から入力された確率的音源符号に対応する確率的音源ベクトルを出力する。また、  
25      ゲイン符号帳 6 0 5 は、符号入力部 6 0 1 から入力されるゲイン符号に対応する適応音源ゲイン及び確率的音源ゲインを読み出し、それぞれ適応音源ゲイン乗算部 6 0 6 及び確率的音源ゲイン乗算部 6 0 7 へ出力する。

そして、適応音源ゲイン乗算部 606 が、適応符号帳 603 から出力された適応音源ベクトルに、ゲイン符号帳 605 から出力された適応音源ゲインを乗算し、確率的音源ゲイン乗算部 607 が、パルス拡散符号帳 604 から出力された確率的音源ベクトルに、ゲイン符号帳 605 で出力された確率的音源ゲインを乗算する。そしてベクトル加算部 608 が、適応音源ゲイン乗算部 606 及び確率的音源ゲイン乗算部 607 それぞれの出力ベクトルを加算して駆動音源ベクトルを生成する。そして、当該駆動音源ベクトルで、合成フィルタ 609 を駆動し、受信したフレーム区間の合成音声を出力する。

このような CELP 方式の音声符号化装置・音声復号化装置において、品質の高い合成音声を得るためには、式 16 の歪み  $ER$  を小さく抑えることが必要になる。そのためには、式 16 の  $ER$  を最小化するように、適応音源符号、確率的音源符号、ゲイン符号の組み合わせを閉ループで特定することが望ましい。しかしながら、式 16 の歪み  $ER$  をクローズドループで特定しようとする演算処理量が大きくなりすぎるため、上記 3 種類の符号はオープンループで特定していくことが一般的である。

具体的には、まず、適応符号帳探索を行う。ここで、適応符号帳探索処理とは、入力音声を逆フィルタに通して得られる予測残差ベクトル中の周期性成分を、過去フレームの駆動音源ベクトルを格納した適応符号帳から出力される適応音源ベクトルによってベクトル量子化する処理である。そして、線形予測残差ベクトル内の周期成分と、近い周期成分を有する適応音源ベクトルのエントリ番号を適応音源符号として特定する。なお、適応符号帳探索によって、同時に、理想適応音源ゲインが暫定的に確定されることになる。

次いで、パルス拡散符号帳探索を行う。パルス拡散符号帳探索は、処理フレームの線形予測残差ベクトルから周期成分を除去した成分、すなわち、線形予測残差ベクトルから適応音源ベクトル成分を差し引いた成分（以下、確率的音源ターゲットと呼ぶこともある）を、パルス拡散符号帳に格納された複数の確率的音源ベクトル候補を用いてベクトル量子化する処理である。そして、この

パルス拡散符号帳探索処理により、確率的音源ターゲットを、もっとも歪み少なく符号化する確率的音源ベクトルのエントリ番号を確率的音源符号として特定する。なお、パルス拡散符号帳探索によって、同時に、理想確率的ゲインも暫定的に確定されることになる。

- 5     そして最後に、ゲイン符号帳探索を行う。ゲイン符号帳探索は、適応符号帳探索時に暫定的に得られた理想適応ゲインと、パルス拡散符号帳探索時に暫定的に得られた理想確率的ゲインとの2要素からなるベクトルを、ゲイン符号帳に格納されたゲイン候補ベクトル（適応音源ゲイン候補と確率的音源ゲイン候補の2要素からなるベクトル候補）で歪みが最小になるように符号化（ベクトル量子化）する処理である。そして、ここで選択されるゲイン候補ベクトルの  
10    エントリ番号がゲイン符号として符号出力部へ出力される。

ここでは、次に、CELP音声符号化装置における上記一般的な符号探索処理のうち、パルス拡散符号帳探索処理（適応音源符号を特定した後に、確率的音源符号を特定する処理）についてさらに詳しく説明を行う。

- 15    説明したように、一般的なCELP符号化装置では、パルス拡散符号帳探索を行う時点では、線形予測符号及び適応音源符号は、既に特定されている。ここで、既に特定されている線形予測符号によって構成される合成フィルタのインパルス応答行列をH、適応音源符号と対応する適応音源ベクトルをp、適応音源符号を特定した時点で同時に求まる理想適応音源ゲイン（暫定値）を $g_a$   
20    とすると、式16の歪みERは、下記式17へと変形される。

$$ER_k = \|v - g_c H c_k\|^2$$

$v$ : 確率的音源ターゲット(ただし、 $v = u - g_a H p$ )

$g_c$ : 確率的音源ゲイン

H: 合成フィルタのインパルス応答行列

- 25     $c_k$ : 確率的音源ベクトル(k: エントリー番号)

ただし、式 17 内のベクトル  $v$  は、フレーム区間内の入力音声信号  $u$ 、合成フィルタのインパルス応答行列  $H$ （既定）、適応音源ベクトル  $p$ （既定）、理想適応音源ゲイン  $g_a$ （暫定値）を用いた、下記式 18 の確率的音源ターゲットである。

$$5 \quad v = u - g_a H p$$

$u$ : 入力音声 (ベクトル)

$g_a$ : 適応音源ゲイン (暫定値)

$H$ : 合成フィルタのインパルス応答行列

$p$ : 適応音源ベクトル

10

式 18

なお、式 16 では確率的音源ベクトルが  $c$  と表現されており、一方、式 17 では確率的音源ベクトルは  $c_k$  と表現がされている。これは、式 16 では確率的音源ベクトルのエン트리番号 ( $k$  のこと) を違いを明示していないことに対して、式 17 ではエン트리番号を明示していることによるものであり、表現上の違いはあるものの意味する対象は同じものである。

したがって、パルス拡散符号帳探索とは、式 17 の歪み  $E R_k$  を最小化するような確率的音源ベクトル  $c_k$  のエン트리番号  $k$  を求める処理である。そして、式 17 の歪み  $E R_k$  を最小化するような確率的音源ベクトル  $c_k$  のエン트리番号  $k$  を特定する際には、確率的音源ゲイン  $g_c$  は任意の値をとりうると仮定できる。したがって、式 17 の歪みを最小化するようなエン트리番号を求める処理は、上記式 10 の分数式  $D_k$  を最大化するような確率的音源ベクトル  $c_k$  のエン트리番号  $k$  を特定する処理に置き換えられる。

25     そして、パルス拡散符号帳探索は、確率的音源ベクトル  $c_k$  のエン트리番号  $k$  ごとに式 10 の分数式  $D_k$  を歪み計算部 505 で計算し、その値を符号特定部 512 へ出力し、符号特定部 512 で、エン트리番号  $k$  ごとの式 10 の値を

大小比較して、その値が最大になるときのエントリ番号  $k$  を確率的音源符号と決定して符号出力部 513 へ出力する、といった 2 段階の処理によって行われることになる。

以下、本実施の形態における音声符号化装置及び音声復号化装置の動作について説明する。

図 11 に示す音声符号化装置におけるパルス拡散符号帳 507 の構成を図 13A に示し、図 12 に示す音声復号化装置におけるパルス拡散符号帳 604 の構成を図 13B に示す。図 13A に示すパルス拡散符号帳 507 と図 13B に示すパルス拡散符号帳 604 を比較した場合、構成上の異なる点は、拡散ボタン格納部に登録している拡散パタンの形状が異なっている点である。

図 13B の音声復号化装置側では、拡散ボタン格納部 4012 には、(1) 多くの確率的音源ターゲットの形状を統計学習し、確率的音源ターゲット中に統計的に高い頻度で含まれる形状の拡散パターン、(2) 無声子音区間や雑音区間を効率的に表現するための乱数的な形状の拡散パターン、(3) 有声定常区間を効率的に表現するためのパルスの形状の拡散パターン、(4) 代数的符号帳から出力される音源ベクトルのエネルギー（非零要素の位置にエネルギーが集中している）を周囲に分散させるような作用を与える形状の拡散パターン、(5) 適当に用意したいくつかの拡散パターン候補について、音声信号を、符号化、復号化、合成音声の視聴評価を繰り返し、品質の高い合成音声を出力しよう  
20 選択した拡散パターン、(6) 音声学的な知見をもとに作成した拡散パタンのうちのいずれかの拡散パターンが各チャンネルあたり 1 種類ずつ登録されている。

一方、図 13A の音声符号化装置側では、拡散ボタン格納部 4012 には、図 13B の音声復号化装置側の拡散ボタン格納部 4012 に登録されている拡散パターンを、1 サンプルおきにゼロに置き換えた拡散パターンが登録されてい  
25 る。

そして、上述のように構成された CELP 音声符号化装置／音声復号化装置では、符号化装置側と復号化装置側で異なる拡散パターンが登録されていること

を意識せずに、上記と同様の方法で、音声信号を符号化・復号化する。

符号化装置では、パルス拡散符号帳を確率的符号帳部に用いた場合の確率的符号帳探索時の前処理演算量を削減することができ ( $H_i = H_t W_i$  及び  $x_{it} = v_t H_i$  の演算量をおよそ半分に削減でき)、復号化装置側では、従来どおりの拡散パターンをパルスベクトルに重畳することで、非零要素位置に集中して  
5 いるエネルギーを周囲に拡散することができ、合成音声の品質を向上することが可能となる。

なお、本実施の形態では、図 1 3 A 及び図 1 3 B に示すように、音声符号化装置側では、音声復号化装置側で用いる拡散パターンを 1 サンプルおきにゼロに  
10 置き換えた拡散パターンを用いる場合について説明したが、音声符号化装置側では、音声復号化装置側で用いる拡散パターの要素を  $N$  ( $N \geq 1$ ) サンプルおきにゼロに置き換えて得られる拡散パターンを用いた場合にも、本実施の形態をそのまま適用することができ、その場合にも同様の作用を得ることができる。

また、本実施の形態では、拡散パターン格納部が、チャンネルあたり 1 種類ずつ  
15 の拡散パターンを登録している場合の実施の形態を説明したが、チャンネルあたり 2 種類以上の拡散パターンが登録されており、それら拡散パターンを選択して用いることを特徴とするパルス拡散符号帳を確率的符号帳部に用いる CELP 音声符号化装置・復号化装置においても本発明を適用することが可能であり、その場合にも同様の作用・効果を得ることができる。

また、本実施の形態では、代数的符号帳部が 3 個の非零要素を含むベクトル  
20 を出力するパルス拡散符号帳を用いた場合について実施の形態を説明したが、代数的符号帳部が出力するベクトル中の非零要素数が  $M$  個 ( $M \geq 1$ ) の場合においても本実施の形態を適用することが可能であり、その場合にも同様の作用・効果を得ることができる。

また、本実施の形態では、少数個の非零要素からなるパルスベクトルを生成  
25 する符号帳として代数的符号帳を用いた場合について説明したが、当該パルスベクトルを生成する符号帳としては、マルチパルス符号帳やレギュラーパルス

符号帳など、その他の符号帳を用いる場合にも本実施の形態を適用することが可能であり、その場合にも同様の作用・効果を得ることができる。

次に、図 1 1 に示す音声符号化装置におけるパルス拡散符号帳の構成を図 1 4 A に示す、図 1 2 に示す音声復号化装置におけるパルス拡散符号帳の構成を  
5 図 1 4 B に示す。

図 1 4 A に示すパルス拡散符号帳と図 1 4 B に示すパルス拡散符号帳の構成を比較した場合、構成上の異なる点は、拡散パターン格納部に登録している拡散パタンの長さが異なっている。図 1 4 B の音声復号化装置側では、拡散パターン格納部 4 0 1 2 には、上述した拡散パターンと同様の拡散パターン、すなわち、  
10 (1) 多くの確率的音源ターゲットの形状を統計学習し、確率的音源ターゲット中に統計的に高い頻度で含まれる形状の拡散パターン、(2) 無声子音区間や雑音区間を効率的に表現するための乱数的な形状の拡散パターン、(3) 有声定常区間を効率的に表現するためのパルスの形状の拡散パターン、(4) 代数的符号帳から出力される音源ベクトルのエネルギー（非零要素の位置にエネルギーが集中している）を周囲に分散させるような作用を与える形状の拡散パターン、  
15 (5) 適当に用意したいくつかの拡散パターン候補について、音声信号を、符号化、復号化、合成音声の視聴評価を繰り返し、品質の高い合成音声を出力しようよう選択した拡散パターン、(6) 音声学的な知見をもとに作成した拡散パタンのうちのいずれかの拡散パターンが各チャンネルあたり 1 種類ずつ登録されている。  
20

一方、図 1 4 A の音声符号化装置側では、拡散パターン格納部 4 0 1 2 には、図 1 4 B の音声復号化装置側の拡散パターン格納部に登録されている拡散パターンを、半分の長さで打ち切った拡散パターンが登録されている。

そして、上述のように構成された CELP 音声符号化装置・復号化装置では、  
25 符号化装置側と復号化装置側で異なる拡散パターンが登録されていることを意識せずに、上述した場合と同様の方法で、音声信号を符号化・復号化する。

符号化装置では、パルス拡散符号帳を確率的符号帳部に用いた場合の確率的



符号帳探索時の前処理演算量を削減することができ ( $H_1 = H_t W_1$  および  $x_{1t} = v_t H_1$  の演算量をおよそ半分に削減でき)、復号化装置側では、従来どおりの拡散パターンを利用することで、合成音声の品質向上を実現することが可能となる。

- 5      なお、本実施の形態では、図 1 4 A 及び図 1 4 B に示すように、音声符号化装置側では、音声復号化装置側で用いる拡散パターンを半分の長さで打ち切った拡散パターンを用いる場合について説明したが、音声符号化装置側では、音声符号化装置側で用いる拡散パターンを、さらに短い長さ  $N$  ( $N \geq 1$ ) で打ち切った場合には、確率的符号帳探索時の前処理演算量をさらに削減することが可能になるといった作用が得られる。ただしここで、音声符号化装置側で用いる拡散
- 10      パターンを長さ 1 で打ち切る場合は、拡散パターンを用いない音声符号化装置に相当する (音声復号化装置には拡散パターンが適用されている)。

- また、本実施の形態では、拡散パターン格納部が、チャンネルあたり 1 種類ずつの拡散パターンを登録している場合を説明したが、チャンネルあたり 2 種類以上の
- 15      拡散パターンが登録されており、それら拡散パターンを選択して用いることを特徴とするパルス拡散符号帳を確率的符号帳部に用いる音声符号化装置／音声復号化装置においても本実施の形態を適用することが可能であり、その場合にも同様の作用・効果を得ることができる。

- また、本実施の形態では、代数的符号帳部が 3 個の非零要素を含むベクトル
- 20      を出力するパルス拡散符号帳を用いた場合について実施の形態を説明したが、代数的符号帳部が出力するベクトル中の非零要素数が  $M$  個 ( $M \geq 1$ ) の場合においても本実施の形態を適用することが可能であり、その場合にも同様の作用・効果を得ることができる。

- また、本実施の形態では、音声符号化装置側では、音声復号化装置側で用い
- 25      る拡散パターンを半分の長さで打ち切った拡散パターンを用いる場合について説明したが、音声符号化装置側では、音声復号化装置側で用いる拡散パターンを長さ  $N$  ( $N \geq 1$ ) で打ち切り、さらに、打ち切り後の拡散パターンを  $M$  ( $M \geq 1$ )

サンプルおきにゼロに置き換えることも可能であり、その場合には、符号探索演算量をさらに低減することが可能になる。

- このように本実施の形態によれば、パルス拡散符号帳を確率的符号帳部に用
- 
- いるC E L P方式の音声符号化装置と復号化装置、及び音声符号化復号化シス
- 5 テムにおいて、学習によって獲得された確率的音源ターゲット中に頻繁に含まれる固定波形を拡散パターンとして登録しておき、当該拡散パターンをパルスベクトルに重畳する（反映させる）ことで、確率的音源ターゲットにより近い確率的音源ベクトルを利用することができるため、復号化側で合成音声の品質向上を実現でき、さらには、符号化側で、パルス拡散符号帳を確率的符号帳部に用
- 10 いる場合に問題となることがある確率的符号帳探索の演算量を、従来よりも低く抑えることが可能となるという有利な効果が得られる。

なお、少数個の非零要素からなるパルスベクトルを生成する符号帳として、マルチパルス符号帳やレギュラーパルス符号帳など、その他の符号帳を用いた場合にも同様の作用・効果を得ることができる。

- 15 上記実施の形態1～3に係る音声符号化／復号化は、音声符号化装置／音声復号化装置として説明しているが、これらの音声符号化／復号化をソフトウェアとして構成しても良い。例えば、上記音声符号化／復号化のプログラムをROMに格納し、そのプログラムにしたがってCPUの指示により動作させるように構成しても良い。また、プログラム、適応符号帳、及び確率的符号帳（パルス拡散符号帳）をコンピュータで読み取り可能な記憶媒体に格納し、この記憶媒体のプログラム、適応符号帳、及び確率的符号帳（パルス拡散符号帳）をコンピュータのRAMに記録して、プログラムにしたがって動作させるようにしても良い。このような場合においても、上記実施の形態1～3と同様の作用、効果を呈する。さらに、実施の形態1～3におけるプログラムを通信端末でダ
- 20 ウンロードし、その通信端末でプログラムを動作させるようにしても良い。

なお、上記実施の形態1～3については、個々に実施しても良く、組み合わせて実施しても良い。

本明細書は、1999年8月23日出願の特願平11-235050号、1999年8月24日出願の特願平11-236728号、及び1999年9月2日出願の特願平11-248363号に基づく。これらの内容はすべてここに含めておく。

5

#### 産業上の利用可能性

本発明は、デジタル通信システムにおける基地局装置や通信端末装置に適用することができる。

## 請求の範囲

1. 適応符号帳及び確率的符号帳に格納された適応音源及び確率的音源に対して、入力音声から求めたL P C係数を用いてフィルタリングすることにより、合成音を得るL P C合成手段と、前記適応音源及び前記確率的音源のゲインを求め、さらに前記ゲインを用いて得られる前記入力音声と前記合成音との間の符号化歪みを用いて適応音源及び確率的音源の符号を探索するゲイン演算手段と、求められた符号に対応する適応音源及び確率的音源を用いてゲインの予測符号化を行うパラメータ符号化手段と、を具備し、  
 前記パラメータ符号化手段は、以前のサブフレームの状態に応じて前記予測符号化に用いる予測係数を調整する予測係数調整手段を備える音声符号化装置。
2. 前記予測係数調整手段は、以前のサブフレームの状態が極端に大きな値又は極端に小さな値である場合に、その影響を小さくするように前記予測係数を調整する請求項1記載の音声符号化装置。
3. 前記パラメータ符号化手段は、適応音源のゲインのベクトル及び確率的音源のゲインのベクトル、並びに予測係数を調整する係数を含む符号帳を有することを特徴とする請求項1記載の音声符号化装置。
4. 予測符号化において、状態と予測係数との間の積和を求めるときに、その状態に対応した予測係数調整係数を乗じる請求項3記載の音声符号化装置。
5. 前記適応音源及び前記確率的音源、並びに予測係数調整係数を状態毎に対応させて格納する格納手段を具備する請求項1記載の音声符号化装置。
6. 前記格納手段に格納された前記適応音源及び前記確率的音源の状態を更新する際に、前記予測係数調整係数も更新する請求項5記載のベクトル量子化装置。
7. 適応符号帳及び確率的符号帳に格納された適応音源及び確率的音源に対して、入力音声から求めたL P C係数を用いてフィルタリングすることにより、合成音を得るL P C合成手段と、前記適応音源及び前記確率的音源のゲインを

求めるゲイン演算手段と、前記入力音声と前記合成音との間の符号化歪みを用いて求められた適応音源及び確率的音源、並びに前記ゲインのベクトル量子化を行うパラメータ符号化手段と、を具備し、1つのフレームを複数のサブフレームに分解して符号化を行うC E L P型音声符号化装置であって、

- 5    最初のサブフレームの適応符号帳探索の前に、フレームを構成する複数のサブフレームのピッチ分析を行って相関値を求め、前記相関値を用いて最もピッチ周期に近似する値を算出するピッチ分析手段を備える音声符号化装置。
8. 前記ピッチ分析手段において得られた相関値及び最もピッチ周期に近似する値に基づいて複数のサブフレームのラグの探索範囲を決定する探索範囲設定手段を具備する請求項7記載の音声符号化装置。
- 10   9. 探索範囲設定手段は、前記ピッチ分析手段において得られた相関値及び最もピッチ周期に近似する値を用いて探索範囲の中心となる仮ピッチを求める請求項8記載の音声符号化装置。
- 15   10. 探索範囲設定手段は、仮ピッチの周りの指定の範囲にラグの探索区間を設定する請求項9記載の音声符号化装置。
11. 探索範囲設定手段は、ラグが短い候補を少なくしてラグの探索区間を設定する請求項8記載の音声符号化装置。
12. 探索範囲設定手段は、適応符号帳探索の際に、設定された範囲でラグの探索を行う請求項8記載の音声符号化装置。
- 20   13. 音声符号化プログラム；過去に合成した音源信号が格納された適応符号帳；複数の音源ベクトルを格納した確率的符号帳；を格納し、コンピュータにより読み取り可能な記録媒体であって、前記音声符号化プログラムは、  
     前記適応符号帳及び前記確率的符号帳に格納された適応音源及び確率的音源に対して、入力音声から求めたL P C係数を用いてフィルタリングすること  
 25    により、合成音を得る手順と、  
     前記適応音源及び前記確率的音源のゲインを求める手順と、  
     前記入力音声と前記合成音との間の符号化歪みを用いて求められた適応音

源及び確率的音源、並びに前記ゲインのベクトル量子化を行う手順と、  
を含み、

- ベクトル量子化を行う手順において、複数の量子化対象ベクトルと、予測符号化に用いる予測係数との間の符号化歪みに基づいて量子化対象ベクトルを  
5 求める手順と、以前のサブフレームの状態に応じて前記予測係数を調整する手順と、を含む。

1 4. 音声符号化プログラム；過去に合成した音源信号が格納された適応符号帳；複数の音源ベクトルを格納した確率的符号帳；を格納し、コンピュータにより読み取り可能な記録媒体であって、前記音声符号化プログラムは、

- 10 前記適応符号帳及び前記確率的符号帳に格納された適応音源及び確率的音源に対して、入力音声から求めたLPC係数を用いてフィルタリングすることにより、合成音を得る手順と、

前記適応音源及び前記確率的音源のゲインを求める手順と、

- 前記入力音声と前記合成音との間の符号化歪みを用いて求められた適応音  
15 源及び確率的音源、並びに前記ゲインのベクトル量子化を行う手順と、

最初のサブフレームの適応符号帳探索の前に、フレームを構成する複数のサブフレームのピッチ分析を行って相関値を求め、前記相関値を用いて最もピッチ周期に近似する値を算出する手順と、  
を含む。

## 要 約 書

- 量子化対象ベクトルの代表的サンプルが複数格納されたベクトル符号帳 1094を作成しておく。各ベクトルは3つの要素からなり、ACゲイン、SCゲインの対数値に対応する値、SCの予測係数の調整係数である。予測係数格納部1095には、予測符号化を行なうための係数を格納する。この係数はMAの予測係数でACとSCの2種類を予測次数分格納する。パラメータ計算部1091において、入力された、聴感重み付け入力音声、聴感重み付けLPC合成済み適応音源、聴感重み付けLPC合成済み確率的音源、さらに、復号化ベクトル格納部1096に格納された復号化ベクトル(AC、SC、調整係数)、
- 5 予測係数格納部1095に格納された予測係数(AC、SC)から距離計算に必要なパラメータを計算する。
- 10

---

**THIS PAGE BLANK (USPTO)**



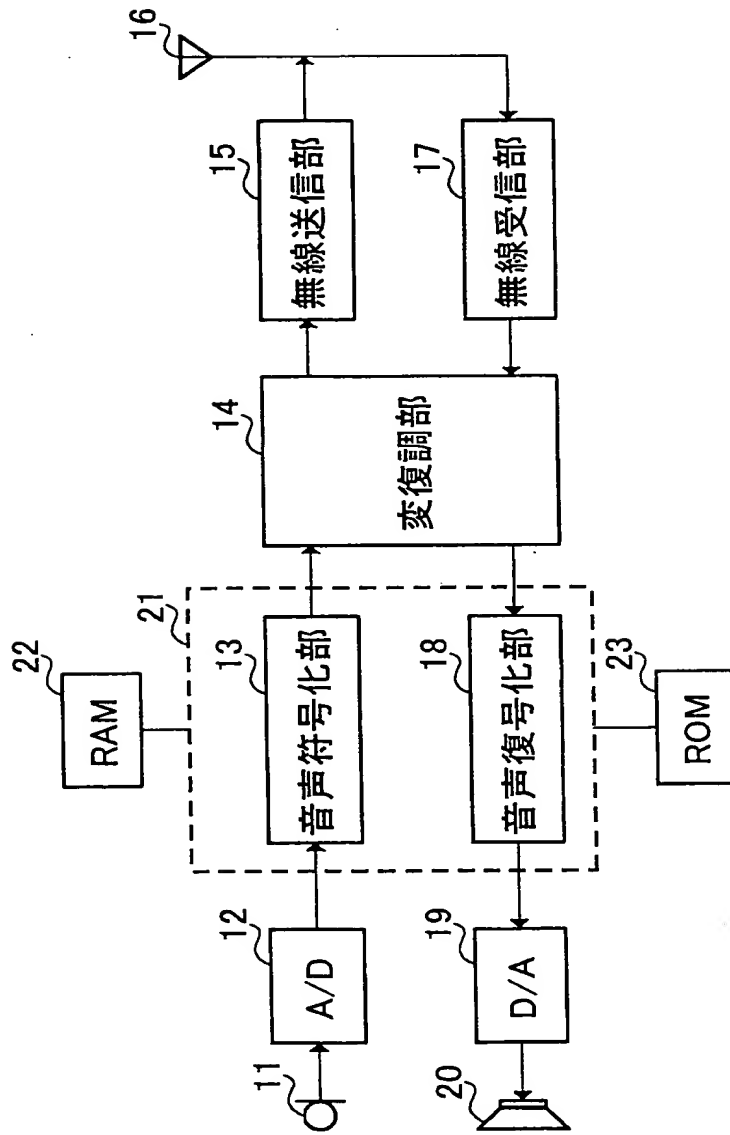


図 1

---

**THIS PAGE BLANK (USPTO)**

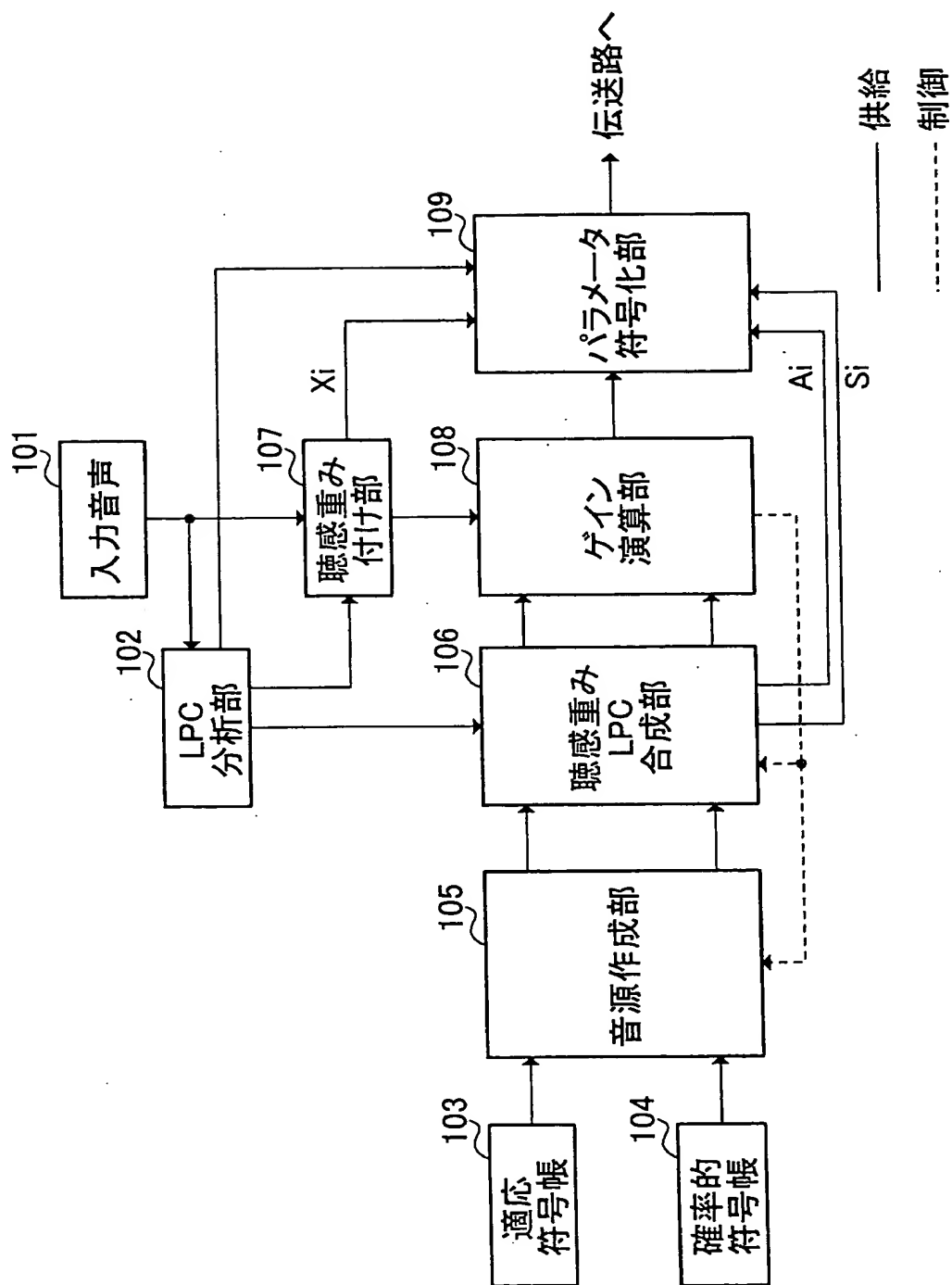
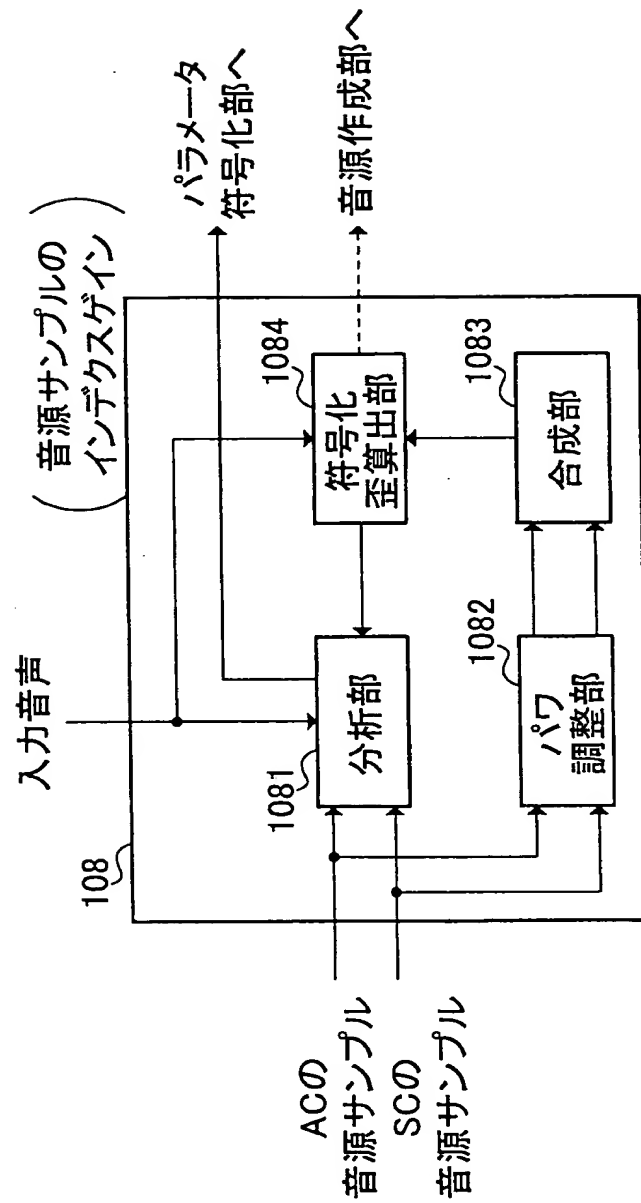


図 2

---

**THIS PAGE BLANK (USPTO)**



3  
X



1 2 3 4 5 6 7 8 9 10 11 12



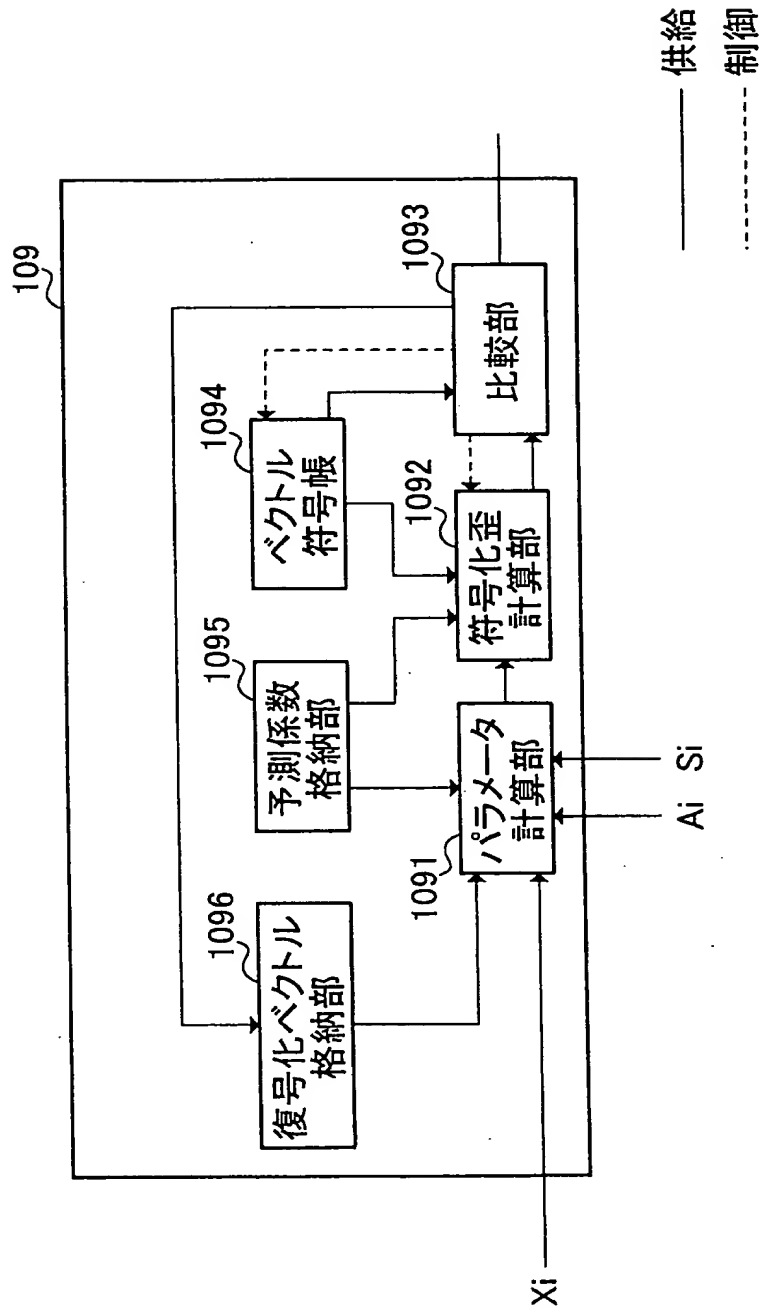


図 4

---

**THIS PAGE BLANK (USPIC)**



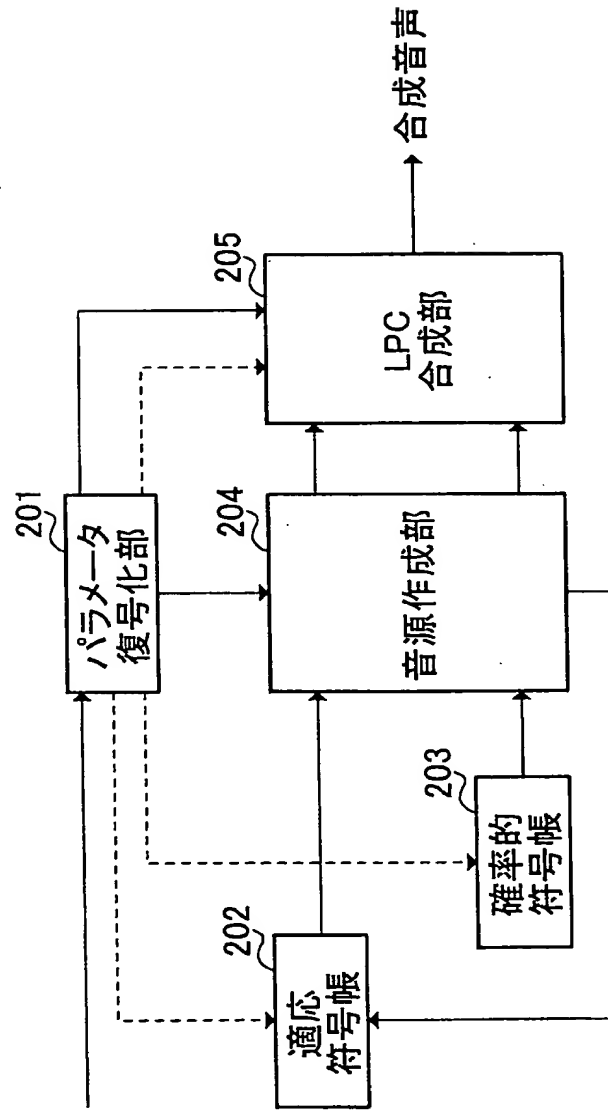


図 5

---

**THIS PAGE BLANK (USPTO)**

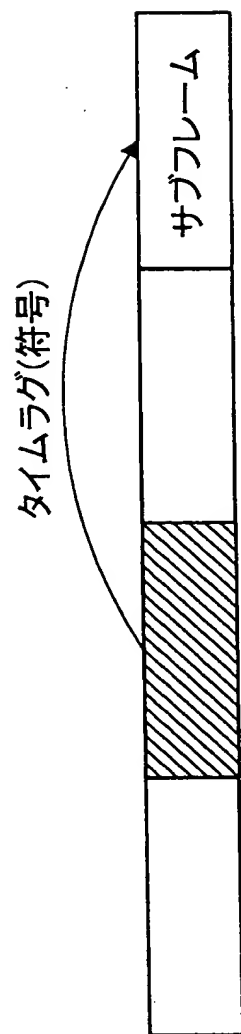
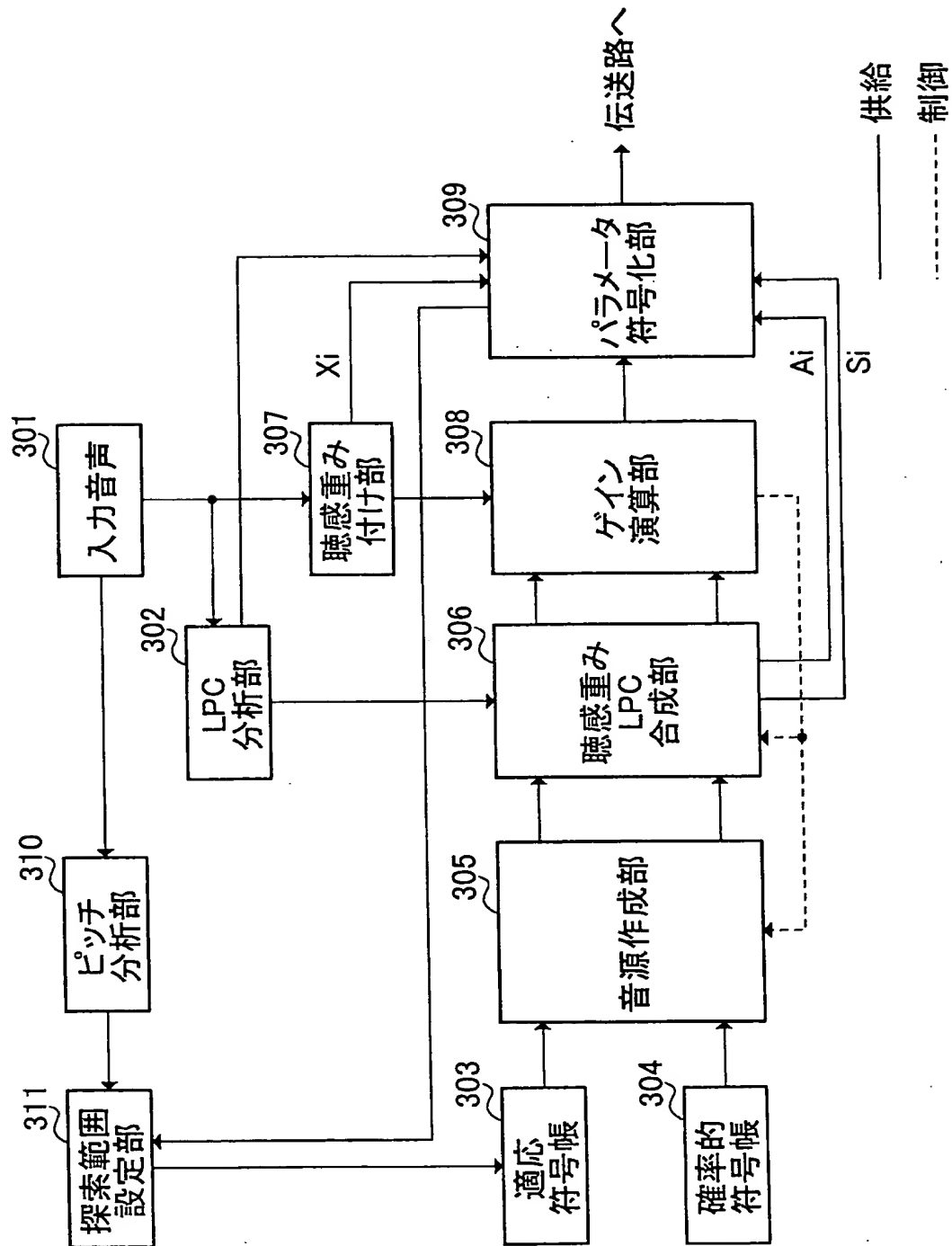


図 6

---

**THIS PAGE BLANK (US)**



7  
✕

---

**THIS PAGE BLANK (000000)**

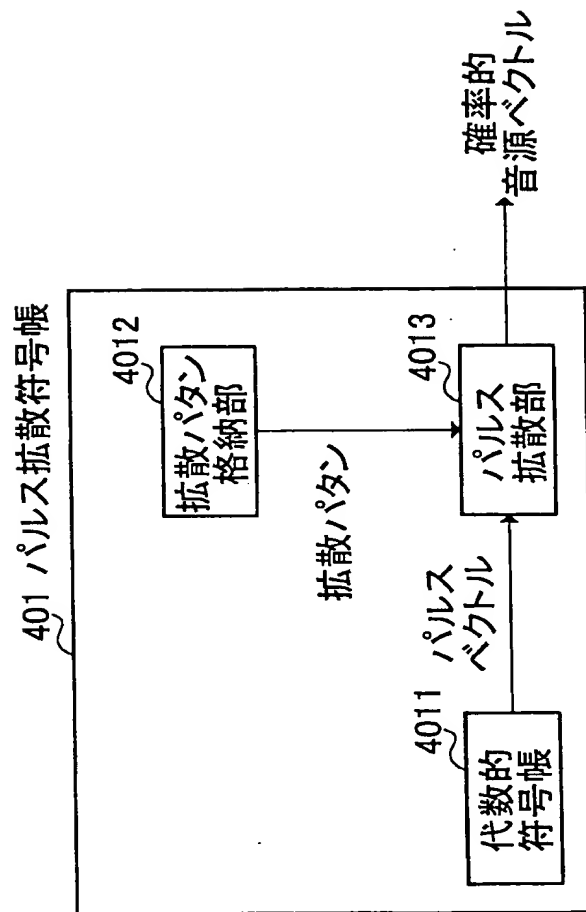


図 8

---

**THIS PAGE BLANK (USPTO)**



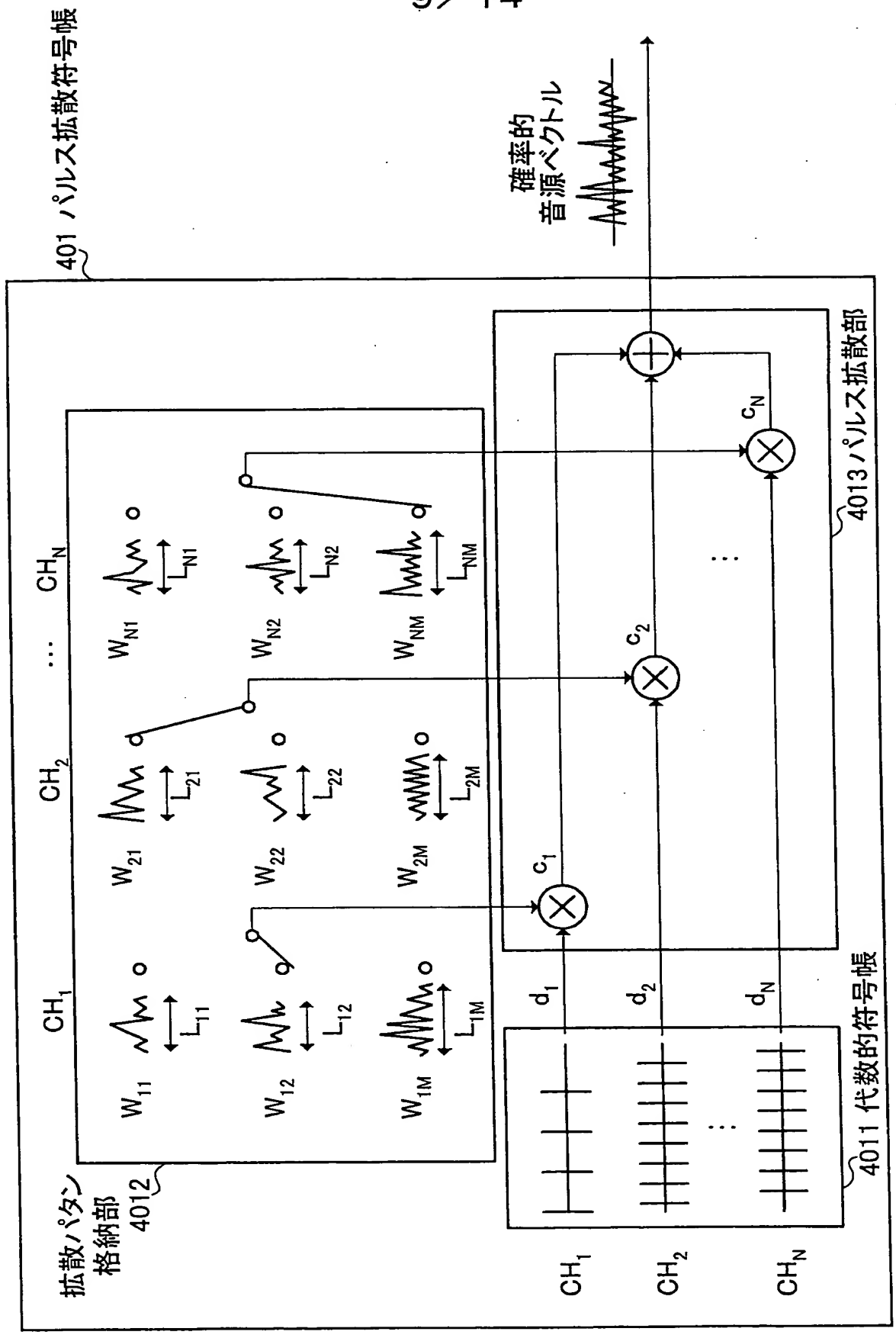


図 9

---

**THIS PAGE BLANK (USPTO)**

401 パルス拡散符号帳

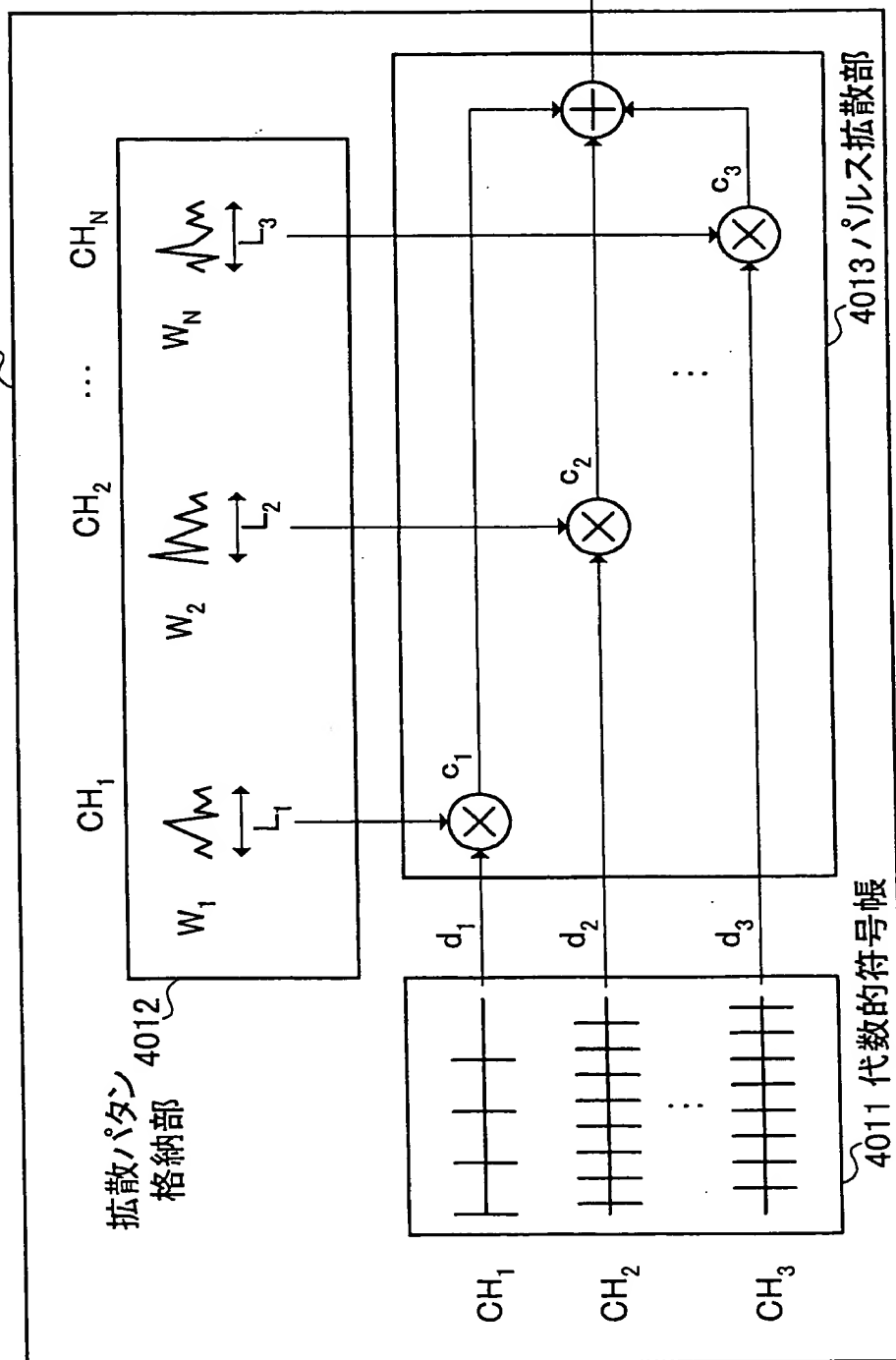


図 10

---

**THIS PAGE BLANK (USPTO)**

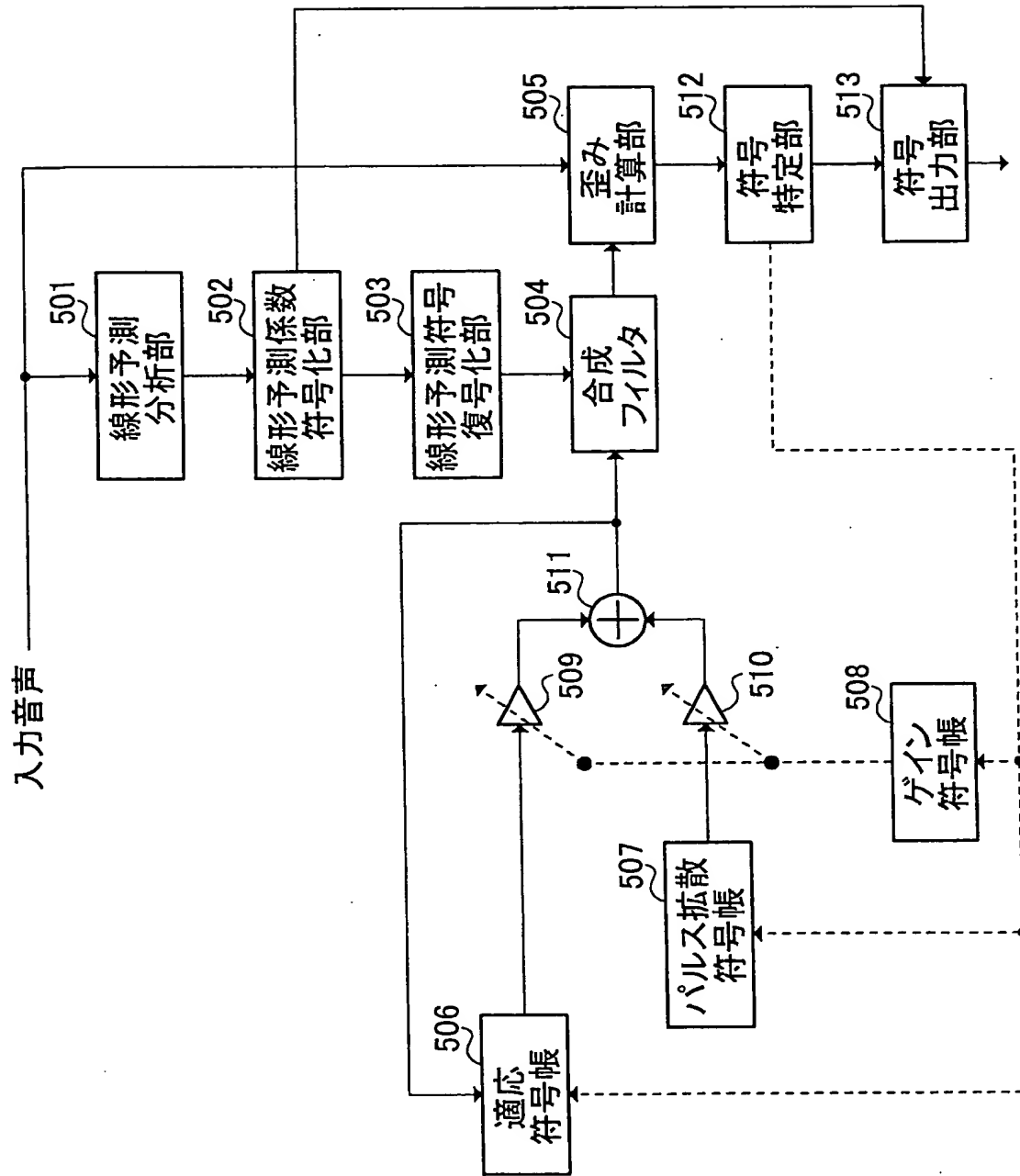


図 11

---

**THIS PAGE BLANK (USPTO)**

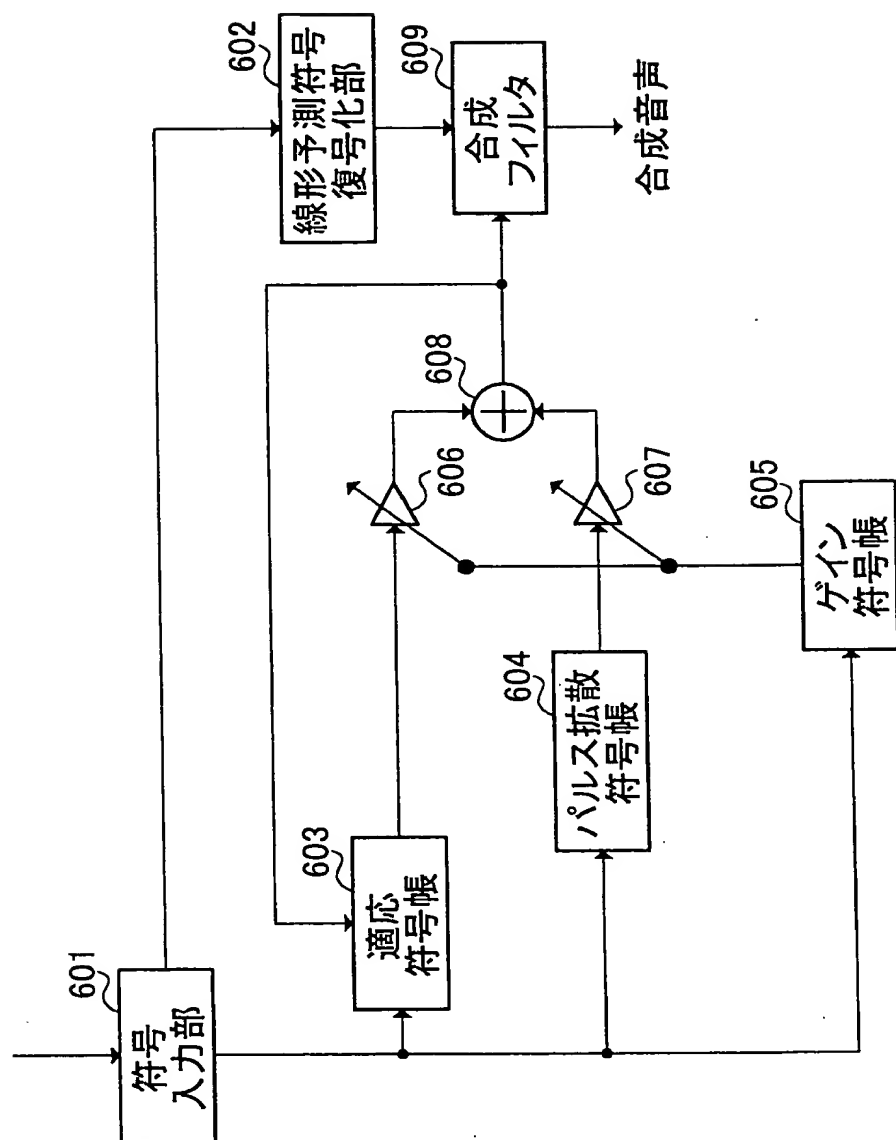


図 12

---

**THIS PAGE BLANK (USPTO)**



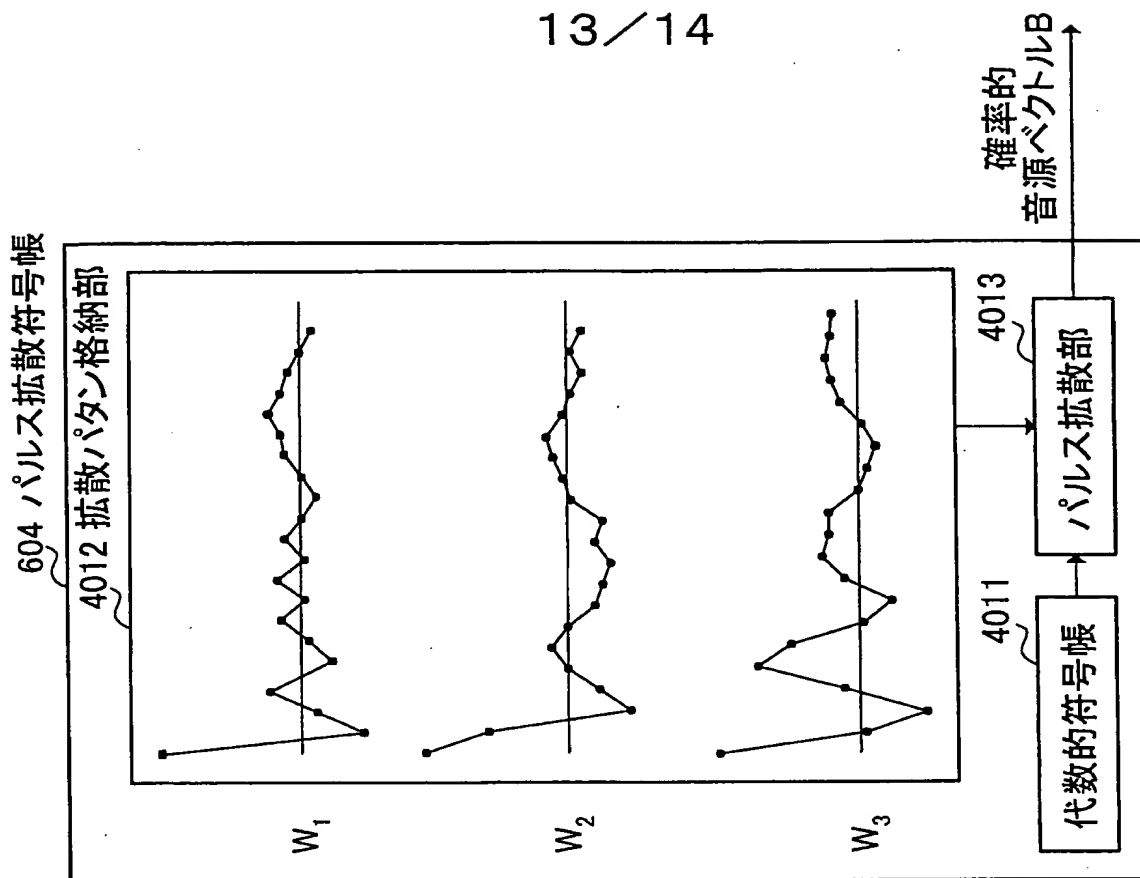


図 13A

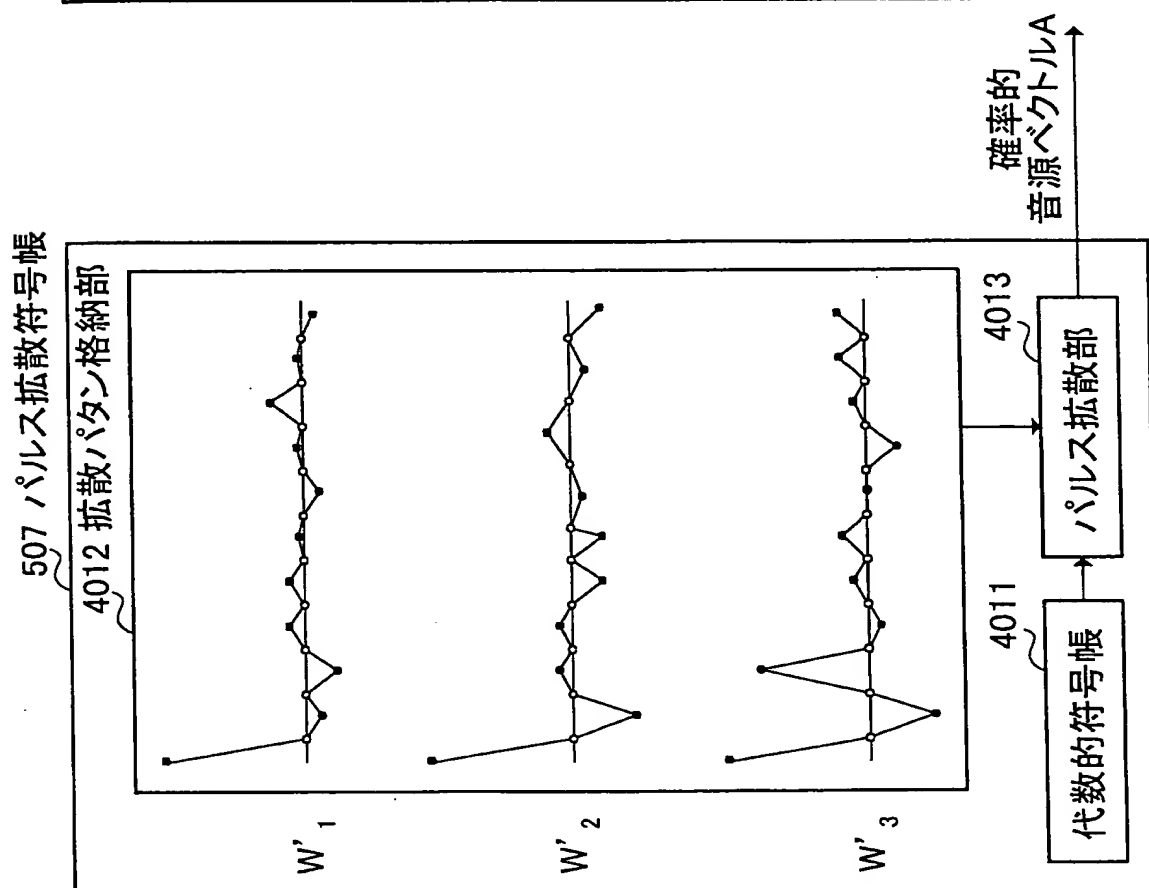


図 13B

---

**THIS PAGE BLANK (USPTO)**

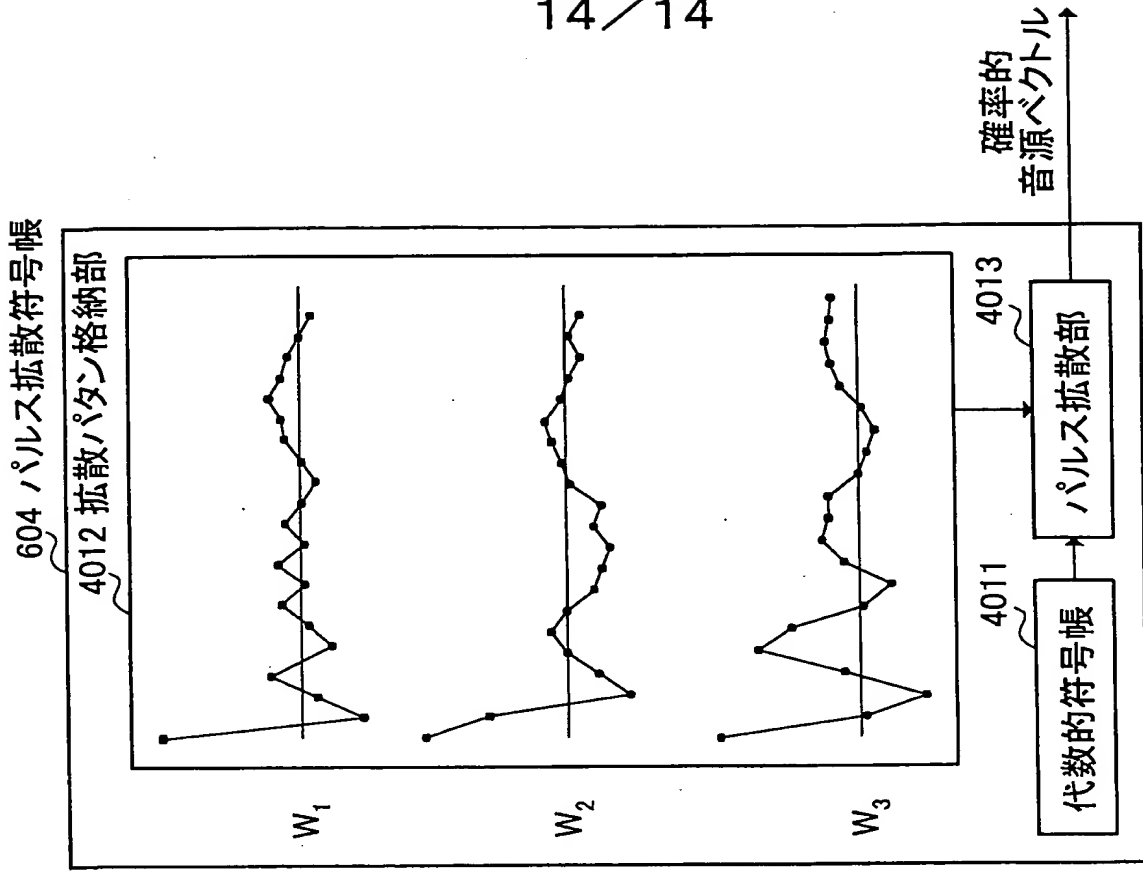


図 14B

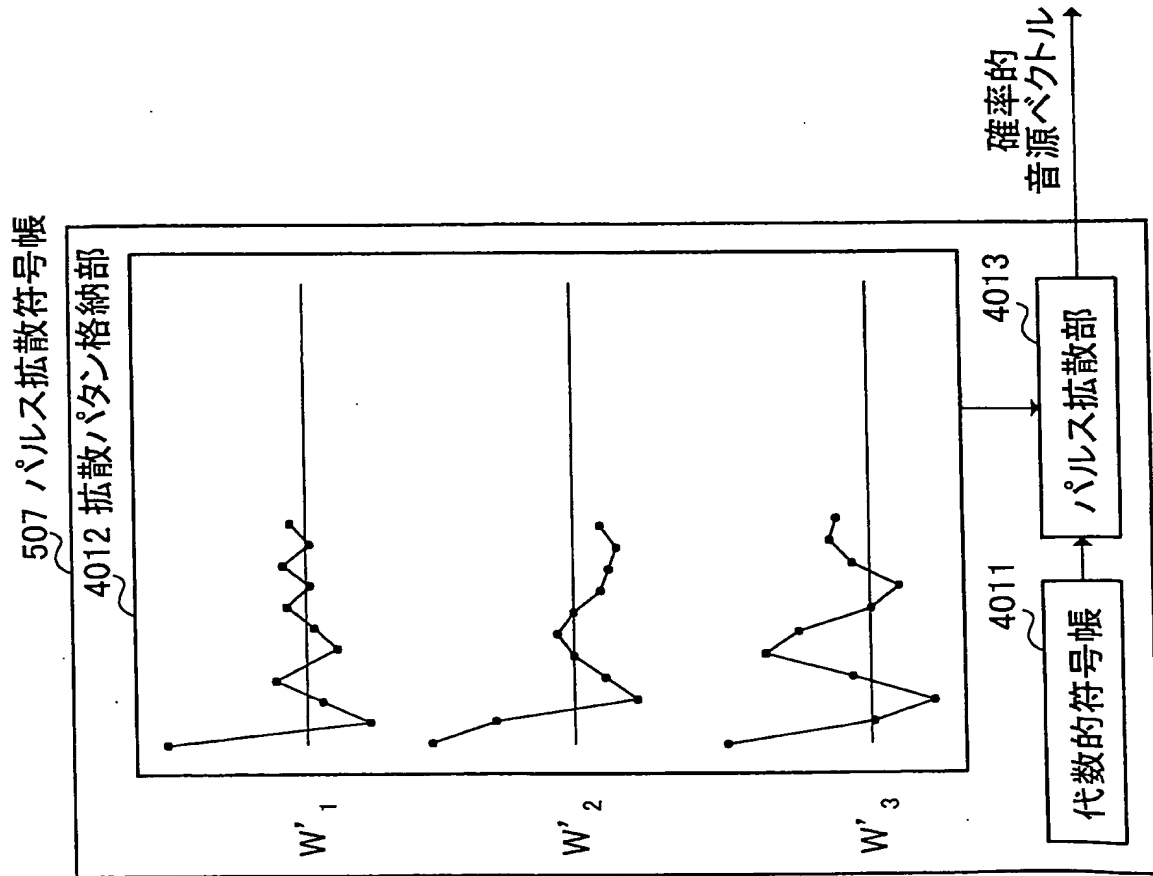


図 14A

---

**THIS PAGE BLANK (COPY)**

## Amended Claims

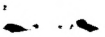
## 1. A speech encoder comprising:

LPC synthesizing means for obtaining a synthesized speech by filtering adaptive excitation vector and stochastic excitation vector stored in an adaptive codebook and stochastic codebook using an LPC coefficients obtained from an input speech;

gain calculating means for calculating gains of said adaptive excitation vector and said stochastic excitation vector and searching code of the adaptive excitation vector and searching code of the stochastic excitation vector using coding distortion between said input speech and said synthesized speech obtained using said gains; and

parameter coding means for performing predictive coding of gains using the adaptive excitation vector and stochastic excitation vector corresponding to the codes obtained, wherein said parameter coding means comprises prediction coefficient adjusting means for adjusting one or more prediction coefficients used for said predictive coding according to one or more states of a previous subframe.

2. The speech encoder according to claim 1, wherein when one or more states of a previous subframes are an extremely large value or an extremely small value, said prediction coefficient adjusting means adjusts said prediction coefficients so as to reduce the influence thereof.



## 補正した請求の範囲

1. 適応符号帳及び確率的符号帳に格納された適応音源及び確率的音源に対して、入力音声から求めたL P C係数を用いてフィルタリングすることにより、合成音を得るL P C合成手段と、前記適応音源及び前記確率的音源のゲインを求め、さらに前記ゲインを用いて得られる前記入力音声と前記合成音との間の符号化歪みを用いて適応音源及び確率的音源の符号を探索するゲイン演算手段と、求められた符号に対応する適応音源及び確率的音源を用いてゲインの予測符号化を行うパラメータ符号化手段と、を具備し、

前記パラメータ符号化手段は、以前のサブフレームの状態に応じて前記予測符号化に用いる予測係数を調整する予測係数調整手段を備える音声符号化装置。

2. 前記予測係数調整手段は、以前のサブフレームの状態が極端に大きな値又は極端に小さな値である場合に、その影響を小さくするように前記予測係数を調整する請求項1記載の音声符号化装置。

3. 前記パラメータ符号化手段は、適応音源のゲインのベクトル及び確率的音源のゲインのベクトル、並びに予測係数を調整する係数を含む符号帳を有することを特徴とする請求項1記載の音声符号化装置。

4. 予測符号化において、状態と予測係数との間の積和を求めるときに、その状態に対応した予測係数調整係数を乗じる請求項3記載の音声符号化装置。

5. 前記適応音源及び前記確率的音源、並びに予測係数調整係数を状態毎に対応させて格納する格納手段を具備する請求項1記載の音声符号化装置。

6. (補正後) 前記格納手段に格納された前記適応音源及び前記確率的音源の状態を更新する際に、前記予測係数調整係数も更新する請求項5記載の音声符号化装置。

7. 適応符号帳及び確率的符号帳に格納された適応音源及び確率的音源に対して、入力音声から求めたL P C係数を用いてフィルタリングすることにより、合成音を得るL P C合成手段と、前記適応音源及び前記確率的音源のゲインを

- 求めるゲイン演算手段と、前記入力音声と前記合成音との間の符号化歪みを用いて求められた適応音源及び確率的音源、並びに前記ゲインのベクトル量子化を行うパラメータ符号化手段と、を具備し、一つのフレームを複数のサブフレームに分解して符号化を行うCELP型音声符号化装置であって、
- 5 最初のサブフレームの適応符号帳探索の前に、フレームを構成する複数のサブフレームのピッチ分析を行って相関値を求め、前記相関値を用いて最もピッチ周期に近似する値を算出するピッチ分析手段を備える音声符号化装置。
  8. 前記ピッチ分析手段において得られた相関値及び最もピッチ周期に近似する値に基づいて複数のサブフレームのラグの探索範囲を決定する探索範囲設定手段を具備する請求項7記載の音声符号化装置。
  - 10 9. 探索範囲設定手段は、前記ピッチ分析手段において得られた相関値及び最もピッチ周期に近似する値を用いて探索範囲の中心となる仮ピッチを求める請求項8記載の音声符号化装置。
  - 15 10. 探索範囲設定手段は、仮ピッチの周りの指定の範囲にラグの探索区間を設定する請求項9記載の音声符号化装置。
  11. 探索範囲設定手段は、ラグが短い候補を少なくしてラグの探索区間を設定する請求項8記載の音声符号化装置。
  12. 探索範囲設定手段は、適応符号帳探索の際に、設定された範囲でラグの探索を行う請求項8記載の音声符号化装置。
  - 20 13. 音声符号化プログラム；過去に合成した音源信号が格納された適応符号帳；複数の音源ベクトルを格納した確率的符号帳；を格納し、コンピュータにより読み取り可能な記録媒体であって、前記音声符号化プログラムは、  
前記適応符号帳及び前記確率的符号帳に格納された適応音源及び確率的音源に対して、入力音声から求めたLPC係数を用いてフィルタリングすること  
25 により、合成音を得る手順と、  
前記適応音源及び前記確率的音源のゲインを求める手順と、  
前記入力音声と前記合成音との間の符号化歪みを用いて求められた適応音



源及び確率的音源、並びに前記ゲインのベクトル量子化を行う手順と、  
を含み、

- ベクトル量子化を行う手順において、複数の量子化対象ベクトルと、予測符号化に用いる予測係数との間の符号化歪みに基づいて量子化対象ベクトルを  
5 求める手順と、以前のサブフレームの状態に応じて前記予測係数を調整する手順と、を含む。

1 4. 音声符号化プログラム；過去に合成した音源信号が格納された適応符号帳；複数の音源ベクトルを格納した確率的符号帳；を格納し、コンピュータにより読み取り可能な記録媒体であって、前記音声符号化プログラムは、

- 10 前記適応符号帳及び前記確率的符号帳に格納された適応音源及び確率的音源に対して、入力音声から求めたLPC係数を用いてフィルタリングすることにより、合成音を得る手順と、

前記適応音源及び前記確率的音源のゲインを求める手順と、

- 前記入力音声と前記合成音との間の符号化歪みを用いて求められた適応音  
15 源及び確率的音源、並びに前記ゲインのベクトル量子化を行う手順と、

最初のサブフレームの適応符号帳探索の前に、フレームを構成する複数のサブフレームのピッチ分析を行って相関値を求め、前記相関値を用いて最もピッチ周期に近似する値を算出する手順と、  
を含む。

- 20 1 5. (追加) 少なくとも一つの非零要素（非零要素以外の要素はゼロの値を持つ）を含むベクトルと、拡散パターンと呼ばれる固定波形とを重畳してベクトルを生成するパルス拡散符号帳を備え、前記パルス拡散符号帳が、音声復号化装置側のパルス拡散符号帳の構成と異なる構成を有する音声符号化装置。

- 1 6. (追加) パルス拡散符号帳の構成部位である拡散パターン格納部が、音声  
25 復号化装置側の拡散パターン格納部が格納している拡散パターンと異なる拡散パターンを格納している請求項 1 5 記載の音声符号化装置。

- 1 7. (追加) 拡散パターン格納部が、音声復号化装置側の拡散パターン格納部が

格納している拡散パターンを簡素化して選られる得られる拡散パターンを格納している請求項 16 記載の音声符号化装置。

18. (追加) 拡散パターン格納部が、音声復号化装置側の拡散パターン格納部が格納している拡散パターの構成要素を、所定の間隔ごとにゼロに置き換えて得られる拡散パターンを格納している請求項 16 記載の音声符号化装置。
19. (追加) 拡散パターン格納部が、音声復号化装置側の拡散パターン格納部が格納している拡散パターの構成要素を、N サンプル (N は自然数) ごとにゼロに置き換えて得られる拡散パターンを格納している請求項 16 記載の音声符号化装置。
20. 20. (追加) 拡散パターン格納部が、音声復号化装置側の拡散パターン格納部が格納している拡散パターの構成要素を、1 サンプルごとにゼロに置き換えて得られる拡散パターンを格納している請求項 19 記載の音声符号化装置。
21. (追加) 拡散パターン格納部が、音声復号化装置側の拡散パターン格納部が格納している拡散パターの構成要素を、適当な長さで打ち切って得られる拡散パターンを格納している請求項 16 記載の音声符号化装置。
22. (追加) 拡散パターン格納部が、音声復号化装置側の拡散パターン格納部が格納している拡散パターの構成要素を、N サンプル (N は自然数) の長さで打ち切って得られる拡散パターンを格納していることを特徴とする請求項 16 記載の音声符号化装置。
20. 23. (追加) 拡散パターン格納部が、音声復号化装置側の拡散パターン格納部が格納している拡散パターの構成要素を、半分の長さに打ち切って得られる拡散パターンを格納している請求項 16 記載の音声符号化装置。
24. (追加) 請求項 15 音声符号化装置で生成された音声符号を有する音声信号を復号化する音声復号化装置。
25. 25. (追加) 請求項 15 音声符号化装置を実現するソフトウェアプログラムを記述した信号処理用プロセッサ。
26. (追加) 請求項 24 記載の音声復号化装置を実現するソフトウェアプロ

グラムを記述した信号処理用プロセッサ。

27. (追加) 音声符号化装置側が有するパルス拡散符号帳の構成と、音声復号化装置側が有するパルス拡散符号帳の構成とが異なる音声符号化復号化システム。

5 28. (追加) 音声符号化装置側が有するパルス拡散符号帳の構成と、音声符号化装置側が有するパルス拡散符号帳の構成との違いが、それぞれのパルス拡散符号帳に備えられた拡散パタンの形状である請求項27記載の音声符号化復号化システム。

29. (追加) 音声符号化装置側の拡散パタンの形状が、音声復号化装置側の  
10 拡散パタンの形状を簡素化したのものである請求項28記載の音声符号化復号化システム。

30. (追加) 音声符号化装置側の拡散パタンの形状が、音声復号化装置側の拡散パタンの構成要素を、適当な間隔ごとにゼロに置き換えて得られる形状である請求項27記載の音声符号化復号化システム。

15 31. (追加) 音声符号化装置側の拡散パタンの形状が、音声復号化装置側の拡散パタンの構成要素を、Nサンプル(Nは自然数)ごとにゼロに置き換えて得られる形状である請求項27記載の音声符号化復号化システム。

32. (追加) 音声符号化装置側の拡散パタンの形状が、音声復号化装置側の拡散パタンの構成要素を、1サンプルごとにゼロに置き換えて得られる形状である  
20 請求項31記載の音声符号化復号化システム。

33. (追加) 音声符号化装置側の拡散パタンの形状が、音声復号化装置側の拡散パタンの構成要素を、適当な長さで打ち切って得られる形状である請求項27記載の音声符号化復号化システム。

34. (追加) 音声符号化装置側の拡散パタンの形状が、音声符号化装置側の  
25 拡散パタンの構成要素を、Nサンプル(Nは自然数)の長さで打ち切って得られる形状である請求項27記載の音声符号化復号化システム。

35. (追加) 音声符号化装置側の拡散パタンの形状が、音声復号化装置側の

拡散パタンの構成要素を、半分の長さに打ち切って得られる形状である請求項 27 記載の音声符号化復号化システム。

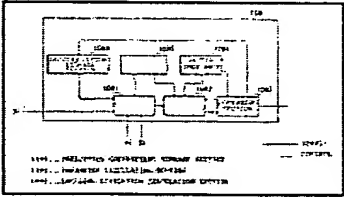
36. (追加) 請求項 25 記載の信号処理用プロセッサを備える通信用基地局

---

5 37. (追加) 請求項 25 記載の信号処理用プロセッサを備える通信用端末。

38. (追加) 請求項 36 記載の通信用基地局及び請求項 37 記載の通信端末を無線ネットワークでつないだ無線通信システム。

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT  
COOPERATION TREATY (PCT)

<p>(51) International Patent Classification <sup>7</sup>: G10L 19/04 // 101:12</p>	A1	<p>(11) Int. Publication WO 01/15144 Number:</p> <p>(43) Int. Publication Date: 01 March 2001 (01.03.2001)</p>
<p>(21) Int. Application Number: PCT/JP00/05621</p> <p>(22) Int. Filing Date: 23 August 2000 (23.08.2000)</p> <p>(30) Priority Data</p> <p>11/235050 Filed on 23 August 1999 JP (23.08.1999)</p> <p>11/236728 Filed on 24 August 1999 JP (24.08.1999)</p> <p>11/248363 Filed on 02 September 1999 JP (02.09.1999)</p> <p>(71) Applicant:</p> <p>MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD. 1006, Oaza Kadoma, Kadoma- shi, Osaka 571-8501; (JP). [JP/JP]. (<i>for all designated states except US</i>)</p> <p>(72) Inventors; and</p> <p>(75) Inventors/Applicants:</p> <p>YASUNAGA, Kazutoshi 3-33-17-305, Sugao, Miyamae-ku, Kawasaki-shi, Kanagawa 216-0015; (JP) [JP/JP]. MORII, Toshiyuki JP JP JP MORII, Toshiyuki 2-3-7-501, Nijigaoka, Asao-ku, Kawasaki-shi, Kanagawa 215- 0015; (JP) [JP/JP].</p> <p>(74) Agent:</p> <p>WASHIDA, Kimihito 5th floor, Shintoshicenter Building, 24-1, Tsurumaki 1-chome, Tama-shi, Tokyo 206-0034; (JP).</p>	<p>(81) Designated States:</p> <p>AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW ; ARIPO patent ( GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW ); Eurasian patent ( AM, AZ, BY, KG, KZ, MD, RU, TJ, TM ); European patent ( AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE ); OAPI patent ( BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG )</p> <p>Published</p> <p><i>With international search report.</i></p> <p><i>With amended claims.</i></p>	
<p>(54) Title: VOICE ENCODER AND VOICE ENCODING METHOD</p> <p>(57) Abstract</p> <p>A vector code book (1094) where representative samples of vectors to be quantized are stored is created. Each vector is made up of three elements: an AC gain, a value corresponding the logarithm of an SC gain, and an adjustment coefficient of the prediction coefficient of SC. Coefficients for predictive coding are stored in a prediction coefficient storage section (1095). The coefficients are the prediction coefficients of MA, and two kinds of coefficients, AC and SC for the order of prediction are stored. A parameter calculating section (1091) calculates a parameter necessary for distance calculation from an auditory sensation weighting input voice, an adaptive sound source subjected to auditory weighting LPC synthesis, a probabilistic sound source subjected to auditory sensation weighting LPC synthesis, a decoded vector (AC, SC, adjustment coefficient) stored in a decoded vector storage section (1096), and the prediction coefficients (AC, SC) stored in the prediction coefficient storage section (1095).</p> <div data-bbox="1003 1438 1344 1633">  </div>		

---

**THIS PAGE BLANK (USPTO)**

(19) 世界知的所有権機関  
国際事務局(43) 国際公開日  
2001年3月1日 (01.03.2001)

PCT

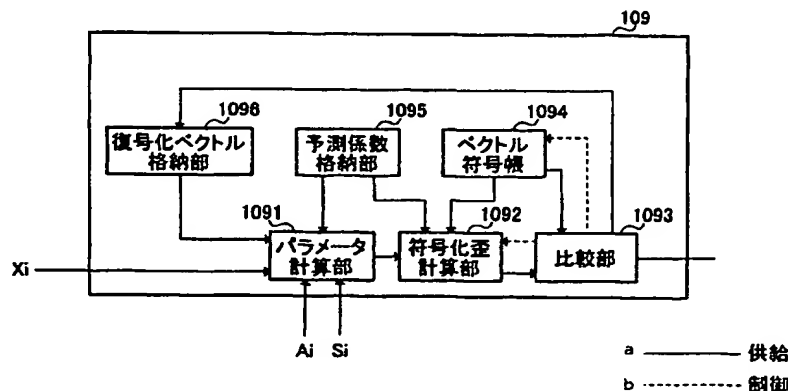
(10) 国際公開番号  
WO 01/15144 A1

- (51) 国際特許分類: G10L 19/04 // 101:12 (72) 発明者; および  
(75) 発明者/出願人 (米国についてのみ): 安永和敏 (YA-SUNAGA, Toshiyuki) [JP/JP]; 〒216-0015 神奈川県川崎市宮前区菅生3-33-17-305 Kanagawa (JP). 森井利幸 (MORII, Kazutoshi) [JP/JP]; 〒215-0015 神奈川県川崎市麻生区虹ヶ丘2-3-7-501 Kanagawa (JP).
- (21) 国際出願番号: PCT/JP00/05621
- (22) 国際出願日: 2000年8月23日 (23.08.2000)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:  
特願平11/235050 1999年8月23日 (23.08.1999) JP  
特願平11/236728 1999年8月24日 (24.08.1999) JP  
特願平11/248363 1999年9月2日 (02.09.1999) JP
- (74) 代理人: 鷺田公一 (WASHIDA, Kimihito); 〒206-0034 東京都多摩市鶴牧1丁目24-1 新都市センタービル5階 Tokyo (JP).
- (81) 指定国 (国内): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (71) 出願人 (米国を除く全ての指定国について): 松下電器産業株式会社 (MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD.) [JP/JP]; 〒571-8501 大阪府門真市大字門真1006番地 Osaka (JP).

[続葉有]

(54) Title: VOICE ENCODER AND VOICE ENCODING METHOD

(54) 発明の名称: 音声符号化装置及び音声符号化方法



1096...DECODED VECTOR STORAGE SECTION

1092...ENCODING DISTORTION CALCULATION SECTION

1095...PREDICTION COEFFICIENT STORAGE SECTION

1093...COMPARING SECTION

1094...VECTOR CODE BOOK

a...SUPPLY

1091...PARAMETER CALCULATING SECTION

b...CONTROL

(57) Abstract: A vector code book (1094) where representative samples of vectors to be quantized are stored is created. Each vector is made up of three elements: an AC gain, a value corresponding the logarithm of an SC gain, and an adjustment coefficient of the prediction coefficient of SC. Coefficients for predictive coding are stored in a prediction coefficient storage section (1095). The coefficients are the prediction coefficients of MA, and two kinds of coefficients, AC and SC for the order of prediction are stored. A parameter calculating section (1091) calculates a parameter necessary for distance calculation from an auditory sensation weighting input voice, an adaptive sound source subjected to auditory weighting LPC synthesis, a probabilistic sound source subjected to auditory sensation weighting LPC synthesis, a decoded vector (AC, SC, adjustment coefficient) stored in a decoded vector storage section (1096), and the prediction coefficients (AC, SC) stored in the prediction coefficient storage section (1095).

[続葉有]

WO 01/15144 A1



(84) 指定国 (広域): ARIPO 特許 (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), ユーラシア特許 (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ特許 (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI 特許 (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

— 補正書

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

添付公開書類:

— 国際調査報告書

(57) 要約:

量子化対象ベクトルの代表的サンプルが複数格納されたベクトル符号帳 1094を作成しておく。各ベクトルは3つの要素からなり、ACゲイン、SCゲインの対数値に対応する値、SCの予測係数の調整係数である。予測係数格納部1095には、予測符号化を行なうための係数を格納する。この係数はMAの予測係数でACとSCの2種類を予測次数分格納する。パラメータ計算部1091において、入力された、聴感重み付け入力音声、聴感重み付けLPC合成済み適応音源、聴感重み付けLPC合成済み確率的音源、さらに、復号化ベクトル格納部1096に格納された復号化ベクトル(AC、SC、調整係数)、予測係数格納部1095に格納された予測係数(AC、SC)から距離計算に必要なパラメータを計算する。



## 明 細 書

## 音声符号化装置及び音声符号化方法

## 5 技術分野

本発明は、ディジタル通信システムにおいて使用される音声符号化装置および音声符号化方法に関する。

## 背景技術

- 10 携帯電話などのディジタル移動通信の分野では、加入者の増加に対処するために低ビットレートの音声の圧縮符号化法が求められており、各研究機関において研究開発が進んでいる。

- 日本国内においては、モトローラ社が開発したビットレート 11.2 kbps の VSELP という符号化法がディジタル携帯電話用の標準符号化方式として採用され、同方式を搭載したディジタル携帯電話は 1994 年秋から国内において発売されている。

- また、NTT 移動通信網株式会社の開発したビットレート 5.6 kbps の PSI-CELP という符号化方式が現在製品化されている。これらの方式はいずれも CELP (Code Excited Linear Prediction: M.R.Schroeder "20 High Quality Speech at Low Bit Rates" Proc.ICASSP'85 pp.937-940 に記載されている) という方式を改良したものである。

- この CELP 方式は、音声を音源情報と声道情報とに分離し、音源情報については符号帳に格納された複数の音源サンプルのインデクスによって符号化し、声道情報については LPC (線形予測係数) を符号化すること及び  
25 音源情報符号化の際に声道情報を加味して入力音声とを比較することを行う方法 (A-b-S: Analysis by Synthesis) を採用していることに特徴がある。

このCELP方式においては、まず、入力された音声データ（入力音声）に対して自己相関分析とLPC分析を行ってLPC係数を得て、得られたLPC係数の符号化を行ってLPC符号を得る。さらに、得られたLPC符号を復号

化して復号化LPC係数を得る。一方、入力音声は、LPC係数を用いた聴感

5 重み付けフィルタを用いて聴感重み付けされる。

適応符号帳と確率的符号帳に格納された音源サンプル（それぞれ適応コードベクトル（又は適応音源）、確率的コードベクトル（又は、確率的音源）と呼ぶ）のそれぞれのコードベクトルに対して、得られた復号化LPC係数によってフィルタリングを行い、2つの合成音を得る。

10 そして、得られた2つの合成音と、聴感重み付けされた入力音声との関係を分析し、2つの合成音の最適値（最適ゲイン）を求め、求められた最適ゲインによって合成音をパワー調整し、それぞれの合成音を加算して総合合成音を得る。その後、得られた総合合成音と入力音声との間の符号化歪みを求める。このようにして、全ての音源サンプルに対して総合合成音と入力音声との間の符  
15 号化歪みを求め、符号化歪みが最も小さいときの音源サンプルのインデクスを  
求める。

このようにして得られたゲイン及び音源サンプルのインデクスを符号化し、これらの符号化されたゲイン及び音源サンプルをLPC符号と共に伝送路に送る。また、ゲイン符号と音源サンプルのインデクスに対応する2つの音源か  
20 ら実際の音源信号を作成し、それを適応符号帳に格納すると同時に古い音源サ  
ンプルを破棄する。

なお、一般的には、適応符号帳と確率的符号帳に対する音源探索は、分析区間をさらに細かく分けた区間（サブフレームと呼ばれる）で行われる。

ゲインの符号化（ゲイン量子化）は、音源サンプルのインデクスに対応する  
25 2つの合成音を用いてゲインの量子化歪を評価するベクトル量子化（VQ）に  
よって行われる。

このアルゴリズムにおいては、予めパラメータベクトルの代表的サンプル

(コードベクトル) が複数格納されたベクトル符号帳を作成しておく。次いで、聴感重み付けした入力音声と、適応音源及び確率的音源を聴感重み付けL P C合成したものに対して、ベクトル符号帳に格納されたゲインコードベクトルを用いて符号化歪を下記式1により計算する。

$$5 \quad E_n = \sum_{i=0}^I (X_i - g_n \times A_i - h_n \times S_i)^2$$

式1

ここで、

$E_n$  : n番のゲインコードベクトルを用いたときの符号化歪み

$X_i$  : 聴感重み付け音声

10  $A_i$  : 聴感重み付けL P C合成済み適応音源

$S_i$  : 聴感重み付けL P C合成済み確率的音源

$g_n$  : コードベクトルの要素 (適応音源側のゲイン)

$h_n$  : コードベクトルの要素 (確率的音源側のゲイン)

$n$  : コードベクトルの番号

15  $i$  : 音源データのインデクス

$I$  : サブフレーム長 (入力音声の符号化単位)

次いで、ベクトル符号帳を制御することによって各コードベクトルを用いたときの歪 $E_n$ を比較し、最も歪の小さいコードベクトルの番号をベクトルの符号とする。また、ベクトル符号帳に格納された全てのコードベクトルの中で最も歪みが小さくなるコードベクトルの番号を求め、これをベクトルの符号とする。

上記式1は一見して各n毎に多くの計算を必要とするように見えるが、予めiについての積和を計算しておけばよいので、少ない計算量でnの探索を行うことができる。

25 一方、音声復号化装置 (デコーダ) では、伝送されてきたベクトルの符号に基づいてコードベクトルを求めることによって符号化されたデータを復号化

してコードベクトルを得る。

また、上記アルゴリズムを基本として、従来よりさらなる改良がなされてきた。例えば、人間の音圧の聴覚特性が対数であることを利用し、パワを対数化

して量子化し、そのパワで正規化した2つのゲインをVQする。この方法は、

- 5 日本国PDCハーフレートコーデックの標準方式で用いられている方法である。また、ゲインパラメータのフレーム間相関を利用して符号化する方法（予測符号化）がある。この方法は、ITU-T国際標準G. 729で用いられている方法である。しかしながら、これらの改良によっても十分な性能を得ることができていない。

- 10 これまで人間の聴覚特性やフレーム間相関を利用したゲイン情報符号化法が開発され、ある程度効率の良いゲイン情報の符号化が可能になった。特に、予測量子化によって性能は大きく向上したが、その従来法では、状態としての値として以前のサブフレームの値をそのまま用いて予測量子化を行っていた。しかしながら、状態として格納される値の中には、極端に大きな（小さな）値
- 15 をとるものがあり、その値を次のサブフレームに用いると、次のサブフレームの量子化がうまくいかず、局所的異音になる場合がある。

#### 発明の開示

- 本発明の目的は、予測量子化を用いて局所的異音を生じることなく音声符号
- 20 化を行うことができるCELP型音声符号化装置及び方法を提供することである。

本発明の主題は、予測量子化において前のサブフレームでの状態値が極端に大きな値や極端に小さな値である場合に、自動的に予測係数を調整することにより、局所的異音の発生を防止することである。

25

#### 図面の簡単な説明

図1は、本発明の音声符号化装置を備えた無線通信装置の構成を示すブロッ

ク図；

図 2 は、本発明の実施の形態 1 に係る音声符号化装置の構成を示すブロック図；

図 3 は、図 2 に示す音声符号化装置におけるゲイン演算部の構成を示すブロック図；

図 4 は、図 2 に示す音声符号化装置におけるパラメータ符号化部の構成を示すブロック図；

図 5 は、本発明の実施の形態 1 に係る音声符号化装置で符号化された音声データを復号する音声復号化装置の構成を示すブロック図；

図 6 は、適応符号帳探索を説明するための図；

図 7 は、本発明の実施の形態 2 に係る音声符号化装置の構成を示すブロック図；

図 8 は、パルス拡散符号帳を説明するためのブロック図；

図 9 は、パルス拡散符号帳の詳細な構成の一例を示すブロック図；

図 10 は、パルス拡散符号帳の詳細な構成の一例を示すブロック図；

図 11 は、本発明の実施の形態 3 に係る音声符号化装置の構成を示すブロック図；

図 12 は、本発明の実施の形態 3 に係る音声符号化装置で符号化された音声データを復号する音声復号化装置の構成を示すブロック図；

図 13 A は、本発明の実施の形態 3 に係る音声符号化装置で用いるパルス拡散符号帳の一例を示す図；

図 13 B は、本発明の実施の形態 3 に係る音声復号化装置で用いるパルス拡散符号帳の一例を示す図；

図 14 A は、本発明の実施の形態 3 に係る音声符号化装置で用いるパルス拡散符号帳の一例を示す図；並びに

図 14 B は、本発明の実施の形態 3 に係る音声復号化装置で用いるパルス拡散符号帳の一例を示す図である。

発明を実施するための最良の形態

以下、本発明の実施の形態について、添付図面を参照して詳細に説明する。

(実施の形態 1)

- 5 図 1 は、本発明の実施の形態 1 ～ 3 に係る音声符号化装置を備えた無線通信装置の構成を示すブロック図である。

この無線通信装置において、送信側で音声マイクなどの音声入力装置 1 1 によって電氣的アナログ信号に変換され、A/D変換器 1 2 に出力される。アナログ音声信号は、A/D変換器 1 2 によってデジタル音声信号に変換され、  
10 音声符号化部 1 3 に出力される。音声符号化部 1 3 は、デジタル音声信号に対して音声符号化処理を行い、符号化した情報を変復調部 1 4 に出力する。変復調部 1 4 は、符号化された音声信号をデジタル変調して、無線送信部 1 5 に送る。無線送信部 1 5 では、変調後の信号に所定の無線送信処理を施す。この信号は、アンテナ 1 6 を介して送信される。なお、プロセッサ 2 1 は、適宜  
15 RAM 2 2 及び ROM 2 3 に格納されたデータを用いて処理を行う。

一方、無線通信装置の受信側では、アンテナ 1 6 で受信した受信信号は、無線受信部 1 7 で所定の無線受信処理が施され、変復調部 1 4 に送られる。変復調部 1 4 では、受信信号に対して復調処理を行い、復調後の信号を音声復号化部 1 8 に出力する。音声復号化部 1 8 は、復調後の信号に復号処理を行ってデ  
20 ィジタル復号音声信号を得て、そのデジタル復号音声信号を D/A 変換器 1 9 へ出力する。D/A 変換器 1 9 は、音声復号化部 1 8 から出力されたデジタル復号音声信号をアナログ復号音声信号に変換してスピーカなどの音声出力装置 2 0 に出力する。最後に音声出力装置 2 0 が電氣的アナログ復号音声信号を復号音声に変換して出力する。

- 25 ここで、音声符号化部 1 3 及び音声復号化部 1 8 は、RAM 2 2 及び ROM 2 3 に格納された符号帳を用いて DSP などのプロセッサ 2 1 により動作する。また、これらの動作プログラムは、ROM 2 3 に格納されている。

図2は、本発明の実施の形態1に係るC E L P型音声符号化装置の構成を示すブロック図である。この音声符号化装置は、図1に示す音声符号化部13に含まれている。なお、図2に示す適応符号帳103は図1に示すR A M 22に格納されており、図2に示す確率的符号帳104は図1に示すR O M 23に格納されている。

図2に示す音声符号化装置においては、L P C分析部102において、入力された音声データ101に対して自己相関分析及びL P C分析を行ってL P C係数を得る。また、L P C分析部102では、得られたL P C係数の符号化を行ってL P C符号を得る。さらに、L P C分析部102では、得られたL P C符号を復号化して復号化L P C係数を得る。入力された音声データ101は、聴感重み付け部107に送られ、そこで上記L P C係数を用いた聴感重み付けフィルタを用いて聴感重み付けされる。

次に、音源作成部105において、適応符号帳103に格納された音源サンプル（適応コードベクトル又は適応音源）と確率的符号帳104に格納された音源サンプル（確率的コードベクトル又は、確率的音源）を取り出し、それぞれのコードベクトルを聴感重みL P C合成部106へ送る。さらに、聴感重みL P C合成部106において、音源作成部105で得られた2つの音源に対して、L P C分析部102で得られた復号化L P C係数によってフィルタリングを行い、2つの合成音を得る。

なお、聴感重みL P C合成部106においては、L P C係数や高域強調フィルタや長期予測係数（入力音声の長期予測分析を行うことによって得られる）を用いた聴感重み付けフィルターを併用してそれぞれの合成音に対して聴感重み付けL P C合成を行う。

聴感重みL P C合成部106は、2つの合成音をゲイン演算部108に出力する。ゲイン演算部108は、図3に示す構成を有する。ゲイン演算部108においては、聴感重みL P C合成部106で得られた2つの合成音及び聴感重み付けされた入力音声を分析部1081に送り、そこで2つの合成音と入力

音声との関係进行分析し、2つの合成音の最適値（最適ゲイン）を求める。この最適ゲインは、パワ調整部1082に出力される。

パワ調整部1082では、求められた最適ゲインによって2つの合成音をパワ調整する。パワ調整された合成音は、合成部1083に出力されて、そこで  
5 加算されて総合合成音となる。この総合合成音は、符号化歪算出部1084に出力される。符号化歪算出部1084では、得られた総合合成音と入力音声との間の符号化歪みを求める。

符号化歪算出部1084は、音源作成部105を制御して、適応符号帳103及び確率的符号帳104の全ての音源サンプルを出力させ、全ての音源サン  
10 プルに対して総合合成音と入力音声との間の符号化歪みを求め、符号化歪みが最も小さいときの音源サンプルのインデクスを求める。

次に、分析部1081は、音源サンプルのインデクス、そのインデクスに対応する2つの聴感重み付けLPC合成された音源、及び入力音声をパラメータ符号化部109に送る。

15 パラメータ符号化部109では、ゲインの符号化を行うことによってゲイン符号を得、LPC符号、音源サンプルのインデクスをまとめて伝送路へ送る。また、ゲイン符号とインデクスに対応する2つの音源から実際の音源信号を作成し、それを適応符号帳103に格納すると同時に古い音源サンプルを破棄する。なお、一般的には、適応符号帳と確率的符号帳に対する音源探索は、分析  
20 区間をさらに細かく分けた区間（サブフレームと呼ばれる）で行われる。

ここで、上記構成を有する音声符号化装置のパラメータ符号化部109のゲイン符号化の動作について説明する。図4は、本発明の音声符号化装置のパラメータ符号化部の構成を示すブロック図である。

図4において、聴感重み付け入力音声（ $X_i$ ）、聴感重み付けLPC合成済み適応音源（ $A_i$ ）、及び聴感重み付けLPC合成済み確率的音源（ $S_i$ ）がパラメータ計算部1091に送られる。パラメータ計算部1091では、符号化歪計算に必要なパラメータを計算する。パラメータ計算部1091で計算され  
25



たパラメータは、符号化歪計算部 1092 に出力され、そこで符号化歪が計算される。この符号化歪は、比較部 1093 に出力される。比較部 1093 では、符号化歪計算部 1092 及びベクトル符号帳 1094 を制御して、得られた符号化歪から最も適当とされる符号（復号化ベクトル）を求め、この符号を基に  
5 ベクトル符号帳 1094 から得られるコードベクトルを復号化ベクトル格納部 1096 に出力し、復号化ベクトル格納部 1096 を更新する。

予測係数格納部 1095 は、予測符号化に用いる予測係数を格納する。この予測係数はパラメータ計算及び符号化歪計算に用いられるために、パラメータ計算部 1091 及び符号化歪計算部 1092 に出力される。復号化ベクトル格  
10 納部 1096 は、予測符号化のために状態を格納する。この状態は、パラメータ計算に用いられるため、パラメータ計算部 1091 に出力される。ベクトル符号帳 1094 は、コードベクトルを格納する。

次に、本発明に係るゲイン符号化方法のアルゴリズムについて説明する。

予め、量子化対象ベクトルの代表的サンプル（コードベクトル）が複数格納  
15 されたベクトル符号帳 1094 を作成しておく。各ベクトルは、ACゲイン、SCゲインの対数値に対応する値、及びSCの予測係数の調整係数の3つの要素からなる。

この調整係数は、以前のサブフレームの状態に応じて予測係数を調整する係数である。具体的には、この調整係数は、以前のサブフレームの状態が極端に  
20 大きな値又は極端に小さな値である場合に、その影響を小さくするように設定される。この調整係数は、多数のベクトルサンプルを用いた本発明者らが開発した学習アルゴリズムにより求めることが可能である。ここでは、この学習アルゴリズムについての説明は省略する。

例えば、有声音に多くの頻度で用いるコードベクトルは調整係数を大きく設  
25 定する。すなわち、同じ波形が並んでいる場合には、以前のサブフレームの状態の信頼性が高いので調整係数を大きくして、以前のサブフレームの予測係数をそのまま利用できるようにする。これにより、より効率的な予測を行うこと

ができる。

一方、語頭などに使用するあまり使用頻度の少ないコードベクトルは調整係数を小さくする。すなわち、前の波形と全然違う場合には、以前のサブフレー

ムの状態の信頼性が低い（適応符号帳が機能しないと考えられる）ので、調整  
5 係数を小さくして、以前のサブフレームの予測係数の影響を小さくする。これにより、次の予測の弊害を防いで良好な予測符号化を実現することができる。

このように、各コードベクトル（状態）に応じて予測係数を制御することにより、これまでの予測符号化の性能をさらに向上させることができる。

また、予測係数格納部 1 0 9 5 には、予測符号化を行うための予測係数を格  
10 納しておく。この予測係数は MA (moving average) の予測係数で AC と SC の 2 種類を予測次数分格納する。これらの予測係数値は、一般に、予め多くのデータを用いた学習により求めておく。また、復号化ベクトル格納部 1 0 9 6 には、初期値として無音状態を示す値を格納しておく。

次に、符号化方法について詳細に説明する。まず、パラメータ計算部 1 0 9  
15 1 に聴感重み付け入力音声 ( $X_i$ )、聴感重み付け LPC 合成済み適応音源 ( $A_i$ )、聴感重み付け LPC 合成済み確率的音源 ( $S_i$ ) を送り、さらに復号化ベクトル格納部 1 0 9 6 に格納された復号化ベクトル (AC、SC、調整係数)、予測係数格納部 1 0 9 5 に格納された予測係数 (AC、SC) を送る。これらを用いて符号化歪計算に必要なパラメータを計算する。

20 符号化歪計算部 1 0 9 2 における符号化歪計算は、下記式 2 にしたがって行う。

$$E_n = \sum_{i=0}^l (X_i - G_{an} \times A_i - G_{sn} \times S_i)^2$$

式 2

ここで、

25  $G_{an}$ ,  $G_{sn}$  : 復号化ゲイン

$E_n$  : n 番のゲインコードベクトルを用いたときの符号化歪み

$X_i$  : 聴感重み付け音声

$A_i$  : 聴感重み付けL P C合成済み適応音源

$S_i$  : 聴感重み付けL P C合成済み確率的音源

$n$  : コードベクトルの番号

5  $i$  : 音源ベクトルのインデクス

$I$  : サブフレーム長 (入力音声の符号化単位)

この場合、演算量を少なくするために、パラメータ計算部1091では、コードベクトルの番号に依存しない部分の計算を行う。計算しておくものは、上記予測ベクトルと3つの合成音 ( $X_i$ ,  $A_i$ ,  $S_i$ ) 間の相関、パワである。こ

10 の計算は、下記式3にしたがって行う。

$$D_{xx} = \sum_{i=0}^I X_i \times X_i$$

$$D_{xa} = \sum_{i=0}^I X_i \times A_i \times 2$$

$$D_{xs} = \sum_{i=0}^I X_i \times S_i \times 2$$

$$D_{aa} = \sum_{i=0}^I A_i \times A_i$$

$$D_{as} = \sum_{i=0}^I A_i \times S_i \times 2$$

$$D_{ss} = \sum_{i=0}^I S_i \times S_i$$

式3

$D_{xx}$ ,  $D_{xa}$ ,  $D_{xs}$ ,  $D_{aa}$ ,  $D_{as}$ ,  $D_{ss}$  : 合成音間の相関値、パワ

$X_i$  : 聴感重み付け音声

15  $A_i$  : 聴感重み付けL P C合成済み適応音源

$S_i$  : 聴感重み付けL P C合成済み確率的音源

$n$  : コードベクトルの番号

$i$  : 音源ベクトルのインデクス

$I$  : サブフレーム長 (入力音声の符号化単位)

また、パラメータ計算部 1091 では、復号化ベクトル格納部 1096 に格納された過去のコードベクトルと、予測係数格納部 1095 に格納された予測係数を用いて下記式 4 に示す 3 つの予測値を計算しておく。

$$\begin{aligned} P_{ra} &= \sum_{m=0}^M \alpha_m \times S_{am} \\ P_{rs} &= \sum_{m=0}^M \beta_m \times S_{cm} \times S_{sm} \\ P_{sc} &= \sum_{m=0}^M \beta_m \times S_{cm} \end{aligned}$$

5

式 4

ここで、

$P_{ra}$  : 予測値 (AC ゲイン)

$P_{rs}$  : 予測値 (SC ゲイン)

$P_{sc}$  : 予測値 (予測係数)

10  $\alpha_m$  : 予測係数 (AC ゲイン、固定値)

$\beta_m$  : 予測係数 (SC ゲイン、固定値)

$S_{am}$  : 状態 (過去のコードベクトルの要素、AC ゲイン)

$S_{sm}$  : 状態 (過去のコードベクトルの要素、SC ゲイン)

$S_{cm}$  : 状態 (過去のコードベクトルの要素、SC 予測係数調整係数)

15  $m$  : 予測インデクス

$M$  : 予測次数

上記式 4 から分かるように、 $P_{rs}$ 、 $P_{sc}$  については、従来と異なり調整係数が乗算されている。したがって、SC ゲインの予測値及び予測係数については、調整係数により、以前のサブフレームにおける状態の値が極端に大きい小さい場合に、それを緩和する (影響を小さくする) ことができる。すなわち、状態に応じて適応的に SC ゲインの予測値及び予測係数を変化させることが可能となる。

20

次に、符号化歪計算部 1092 において、パラメータ計算部 1091 で計算

した各パラメータ、予測係数格納部 1095 に格納された予測係数、及びベクトル符号帳 1094 に格納されたコードベクトルを用いて、下記式 5 にしたがって符号化歪を算出する。

$$E_n = D_{xx} + (G_{an})^2 \times D_{aa} + (G_{sn})^2 \times D_{ss} - G_{an} \times D_{xa} - G_{sn} \times D_{xs} + G_{an} \times G_{sn} \times D_{as}$$

$$G_{an} = P_{ra} + (1 - P_{ac}) \times C_{an}$$

$$G_{sn} = 10^{\{P_{rs} + (1 - P_{sc}) \times C_{sn}\}}$$

5

式 5

ここで、

$E_n$  :  $n$  番のゲインコードベクトルを用いたときの符号化歪み

$D_{xx}$ ,  $D_{xa}$ ,  $D_{xs}$ ,  $D_{aa}$ ,  $D_{as}$ ,  $D_{ss}$  : 合成音間の相関値、パワ

$G_{an}$ ,  $G_{sn}$  : 復号化ゲイン

10  $P_{ra}$  : 予測値 (ACゲイン)

$P_{rs}$  : 予測値 (SCゲイン)

$P_{ac}$  : 予測係数の和 (固定値)

$P_{sc}$  : 予測係数の和 (上記式 4 で算出)

$C_{an}$ ,  $C_{sn}$ ,  $C_{cn}$  : コードベクトル、 $C_{cn}$  は予測係数調整係数であるがこ

15 こでは使用しない

$n$  : コードベクトルの番号

なお、実際には  $D_{xx}$  はコードベクトルの番号  $n$  に依存しないので、その加算を省略することができる。

次いで、比較部 1093 は、ベクトル符号帳 1094 と符号化歪計算部 1092 を制御し、ベクトル符号帳 1094 に格納された複数のコードベクトルの中で符号化歪計算部 1092 にて算出された符号化歪みの最も小さくなるコードベクトルの番号を求め、これをゲインの符号とする。また、得られたゲインの符号を用いて復号化ベクトル格納部 1096 の内容を更新する。更新は、

20

下記式 6 にしたがって行う。

$$S_{am} = S_{am-1}(M=M \sim 1), S_{a0} = C_a J$$

$$S_{sm} = S_{sm-1}(M=M \sim 1), S_{s0} = C_s J$$

$$S_{cm} = S_{cm-1}(M=M \sim 1), S_{c0} = C_c J$$

5

式 6

ここで、

$S_{am}$ ,  $S_{sm}$ ,  $S_{cm}$  : 状態ベクトル (AC、SC、予測係数調整係数)

$m$  : 予測インデクス

$M$  : 予測次数

10  $J$  : 比較部で求められた符号

式 4 から式 6 までで分かるように、本実施の形態では、復号化ベクトル格納部 1096 で状態ベクトル  $S_{cm}$  を格納しておいて、この予測係数調整係数を用いて予測係数を適応的に制御している。

図 5 は、本発明の実施の形態の音声復号化装置の構成を示すブロック図である。この音声復号化装置は、図 1 に示す音声復号化部 18 に含まれている。なお、図 5 に示す適応符号帳 202 は図 1 に示す RAM 22 に格納されており、図 5 に示す確率的符号帳 203 は図 1 に示す ROM 23 に格納されている。

図 5 に示す音声復号化装置において、パラメータ復号化部 201 は、伝送路から、符号化された音声信号を得ると共に、各音源符号帳 (適応符号帳 202、確率的符号帳 203) の音源サンプルの符号、LPC 符号、及びゲイン符号を得る。そして、LPC 符号から復号化された LPC 係数を得て、ゲイン符号から復号化されたゲインを得る。

そして、音源作成部 204 は、それぞれの音源サンプルに復号化されたゲインを乗じて加算することによって復号化された音源信号を得る。この際、得られた復号化された音源信号を、音源サンプルとして適応符号帳 204 へ格納し、同時に古い音源サンプルを破棄する。そして、LPC 合成部 205 では、復号化された音源信号に復号化された LPC 係数によるフィルタリングを行うこ

とによって、合成音を得る。

また、2つの音源符号帳は、図2に示す音声符号化装置に含まれるもの（図2の参照符号103，104）と同様のものであり、音源サンプルを取り出すためのサンプル番号（適応符号帳への符号と確率的符号帳への符号）は、いず  
5 れもパラメータ復号化部201から供給される。

このように、本実施の形態の音声符号化装置では、各コードベクトルに応じて予測係数を制御することが可能になり、音声の局所的特徴により適応したより効率的な予測や、非定常部における予測の弊害を防ぐことが可能になり、従  
来得られなかった格別の効果を得ることができる。

#### 10 （実施の形態2）

音声符号化装置において、上述したように、ゲイン演算部では、音源作成部から得られた適応符号帳、確率的符号帳の全ての音源について合成音と入力音声との間の比較を行う。このとき、演算量の都合上、通常は2つの音源（適応符号帳と確率的符号帳）はオープンループに探索される。以下、図2を参照し  
15 て説明する。

このオープンループ探索においては、まず、音源作成部105は適応符号帳103からのみ音源候補を次々に選び、聴感重みLPC合成部106を機能させて合成音を得て、ゲイン演算部108へ送り、合成音と入力音声との間の比較を行って最適な適応符号帳103の符号を選択する。

20 次いで、上記適応符号帳103の符号を固定して、適応符号帳103からは同じ音源を選択し、確率的符号帳104からはゲイン演算部108の符号に対応した音源を次々に選択して聴感重みLPC合成部106へ伝送する。ゲイン演算部108で両合成音の和と入力音声との間の比較を行って確率的符号帳104の符号を決定する。

25 このアルゴリズムを用いた場合、全ての符号帳の符号をそれぞれに対して全て探索するよりは符号化性能は若干劣化するが、計算量は大幅に削減される。このため一般にはこのオープンループ探索が用いられる。

ここで、従来のオープンループの音源探索の中で代表的なアルゴリズムについて説明する。ここでは、1つの分析区間（フレーム）に対して2つのサブフレームで構成する場合の音源探索手順について説明する。

- まず、ゲイン演算部108の指示を受けて、音源作成部105は適応符号帳
- 5 103から音源を引出して聴感重みLPC合成部106へ送る。ゲイン演算部108において、合成された音源と第1サブフレームの入力音声との間の比較を繰り返して最適な符号を求める。ここで、適応符号帳の特徴を示す。適応符号帳は過去において合成に使用した音源である。そして、符号は、図6に示すようにタイムラグに対応している。
- 10 次に、適応符号帳103の符号が決まった後に、確率的符号帳の探索を行う。音源作成部105は適応符号帳103の探索で得られた符号の音源とゲイン演算部108で指定された確率的符号帳104の音源とを取り出して聴感重みLPC合成部106へ送る。そして、ゲイン演算部108において、聴感重み付け済みの合成音と聴感重み付け済みの入力音声との間の符号化歪みを計算し、最も適当な（二乗誤差が最小となるもの）確率的音源104の符号を決める。1つの分析区間（サブフレームが2の場合）での音源符号探索の手順を以下に示す。
- 1) 第1サブフレームの適応符号帳の符号を決定
  - 2) 第1サブフレームの確率的符号帳の符号を決定
  - 20 3) パラメータ符号化部109でゲインを符号化し、復号化ゲインで第1サブフレームの音源を作成し、適応符号帳103を更新する。
  - 4) 第2サブフレームの適応符号帳の符号を決定
  - 5) 第2サブフレームの確率的符号帳の符号を決定
  - 6) パラメータ符号化部109でゲインを符号化し、復号化ゲインで第2サブフレームの音源を作成し、適応符号帳103を更新する。
- 25

上記アルゴリズムによって効率よく音源の符号化を行うことができる。しかしながら、最近では、さらなる低ビットレート化を目指し、音源のビット数を



節約する工夫が行われている。特に注目されているのは、適応符号帳のラグに大きな相関があることを利用して、第1サブフレームの符号はそのまま、第2サブフレームの探索範囲を第1サブフレームのラグの近くに狭めて（エントリ数を減らして）ビット数を少なくするというアルゴリズムである。

- 5      このアルゴリズムでは、分析区間（フレーム）の途中から音声に変化する場合や、2つのサブフレームの様子が大きく異なる場合には局所的劣化を引き起こすことが考えられる。

- 10      本実施の形態では、符号化の前に2つのサブフレーム両方についてピッチ分析を行って相関値を算出し、得られた相関値に基づいて2つのサブフレームのラグの探索範囲を決定する探索方法を実現する音声符号化装置を提供する。

- 15      具体的には、本実施の形態の音声符号化装置は、1つのフレームを複数のサブフレームに分解してそれぞれを符号化するCELP型符号化装置において、最初のサブフレームの適応符号帳探索の前に、フレームを構成する複数のサブフレームのピッチ分析を行って相関値を算出するピッチ分析部と、上記ピッチ分析部がフレームを構成する複数のサブフレームの相関値を算出すると共に、その相関値の大小から各サブフレームで最もピッチ周期らしい値（代表ピッチと呼ぶ）を求め、ピッチ分析部にて得られた相関値と代表ピッチとに基づいて複数のサブフレームのラグの探索範囲を決定する探索範囲設定部と、を備えることを特徴としている。そして、この音声符号化装置では、探索範囲設定部において、ピッチ分析部で得た複数のサブフレームの代表ピッチと相関値を利用して探索範囲の中心となる仮のピッチ（仮ピッチと呼ぶ）を求め、探索範囲設定部において、求めた仮ピッチの周りの指定の範囲にラグの探索区間を設定し、ラグの探索区間を設定するときに、仮ピッチの前後に探索範囲を設定する。また、その際に、ラグの短い部分の候補を少なくし、ラグのより長い範囲を広く
- 20      設定し、適応符号帳探索の際に上記探索範囲設定部で設定された範囲でラグの探索を行う。

以下、本実施の形態に係る音声符号化装置について添付図面を用いて詳細に

説明する。ここでは、1フレームは2サブフレームに分割されているものとする。3サブフレーム以上の場合でも同様の手順で符号化を行うことができる。

この音声符号化装置においては、いわゆるデルタラグ方式によるピッチ探索において、分割されたサブフレームについてすべてピッチを求め、ピッチ間で  
5    どの程度の相関があるかどうかを求めて、その相関結果に応じて探索範囲を決定する。

図7は、本発明の実施の形態2に係る音声符号化装置の構成を示すブロック図である。まず、LPC分析部302において、入力された音声データ（入力音声）301に対して自己相関分析とLPC分析を行うことによってLPC係  
10    数を得る。また、LPC分析部302にばいて、得られたLPC係数の符号化を行ってLPC符号を得る。さらに、LPC分析部302において、得られたLPC符号を復号化して復号化LPC係数を得る。

次いで、ピッチ分析部310において、2サブフレーム分の入力音声のピッチ分析を行い、ピッチ候補とパラメータを求める。1サブフレームに対するアル  
15    ゴリズムを以下に示す。相関係数は、下記式7により、2つ求められる。なおこの時、 $C_{pp}$ は $P_{min}$ についてまず求め、あとの $P_{min+1}$ 、 $P_{min+2}$ については、フレーム端の値の足し引きで効率的に計算できる。

$$V_p = \sum_{i=0}^L X_i \times X_{i-P} \quad (P=P_{min} \sim P_{max})$$

$$C_{pp} = \sum_{i=0}^L X_{i-P} \times X_{i-P} \quad (P=P_{min} \sim P_{max})$$

20

式7

ここで、

$X_i$ ,  $X_{i-P}$  : 入力音声

$V_p$  : 自己相関関数

$C_{pp}$  : パワ成分

25     $i$  : 入力音声のサンプル番号

L : サブフレームの長さ

P : ピッチ

$P_{min}$ ,  $P_{max}$  : ピッチの探索を行う最小値と最大値

そして、上記式 7 で求めた自己相関関数とパワー成分はメモリに蓄えておき、

- 5 次の手順で代表ピッチ  $P_1$  を求める。これは  $V_p$  が正で  $V_p \times V_p / C_{pp}$  を最大にするピッチ  $P$  を求める処理となっている。ただし、割り算は一般的に計算量がかかるので、分子と分母を 2 つとも格納し、掛け算に直して効率化を図っている。

- 10 ここでは、入力音声と入力音声からピッチ分過去の適応音源との差分の二乗和が最も小さくなるようなピッチを探す。この処理は  $V_p \times V_p / C_{pp}$  を最大にするピッチ  $P$  を求める処理と等価となる。具体的な処理は以下ようになる。

1) 初期化 ( $P = P_{min}$ ,  $VV = C = 0$ ,  $P_1 = P_{min}$ )

2) もし ( $V_p \times V_p \times C < VV \times C_{pp}$ ) 又は ( $V_p < 0$ ) ならば 4) へ。それ以外なら 3) へ。

- 15 3)  $VV = V_p \times V_p$ ,  $C = C_{pp}$ ,  $P_1 = P$  として 4) へ

4)  $P = P + 1$  とする。この時  $P > P_{max}$  であれば終了、それ以外の場合には 2) へ。

- 20 上記作業を 2 サブフレームのそれぞれについて行い、代表ピッチ  $P_1$ ,  $P_2$  と自己相関係数  $V_{1p}$ ,  $V_{2p}$ , パワー成分  $C_{1pp}$ ,  $C_{2pp}$  ( $P_{min} < p < P_{max}$ ) を求める。

次に、探索範囲設定部 311 で適応符号帳のラグの探索範囲を設定する。まず、その探索範囲の軸となる仮ピッチを求める。仮ピッチはピッチ分析部 310 で求めた代表ピッチとパラメータを用いて行う。

- 25 仮ピッチ  $Q_1$ ,  $Q_2$  は以下の手順で求める。なお、以下の説明においてラグの範囲として定数  $Th$  (具体的には 6 程度が適当である) を用いる。また、相関値は上記式 7 で求めたものを用いる。

まず、 $P_1$  を固定した状態で  $P_1$  の付近 ( $\pm Th$ ) で相関の最も大きい仮ピッ

チ ( $Q_2$ ) を見つける。

1) 初期化 ( $p = P_1 - Th$ ,  $C_{max} = 0$ ,  $Q_1 = P_1$ ,  $Q_2 = P_1$ )

2) もし ( $V_{1p1} \times V_{1p1} / C_{1p1p1} + V_{2p} \times V_{2p} / C_{2pp} < C_{max}$ ) または ( $V_{2p} < 0$ ) ならば4) へ。それ以外なら3) へ。

5) 3)  $C_{max} = V_{1p1} \times V_{1p1} / C_{1p1p1} + V_{2p} \times V_{2p} / C_{2pp}$ ,  $Q_2 = p$  として4) へ

4)  $p = p + 1$  として2) へ。ただし、この時  $p > P_1 + Th$  であれば5) へ。

10) このようにして2) ~ 4) の処理を  $P_1 - Th \sim P_1 + Th$  まで行って、相関の最も大きいもの  $C_{max}$  と仮ピッチ  $Q_2$  を求める。

次に、 $P_2$  を固定した状態で  $P_2$  の付近 ( $\pm Th$ ) で相関の最も大きい仮ピッチ ( $Q_1$ ) を求める。この場合、 $C_{max}$  は初期化しない。 $Q_2$  を求めた際の  $C_{max}$  を含めて相関が最大となる  $Q_1$  を求めることにより、第1, 第2サブフレーム間で最大の相関を持つ  $Q_1$ ,  $Q_2$  を求めることが可能となる。

15) 5) 初期化 ( $p = P_2 - Th$ )

6) もし ( $V_{1p} \times V_{1p} / C_{1pp} + V_{2p2} \times V_{2p2} / C_{2p2p2} < C_{max}$ ) 又は ( $V_{1p} < 0$ ) ならば8) へ。それ以外は7) へ。

7)  $C_{max} = V_{1p} \times V_{1p} / C_{1pp} + V_{2p2} \times V_{2p2} / C_{2p2p2}$ ,  $Q_1 = p$ ,  $Q_2 = P_2$  として8) へ。

20) 8)  $p = p + 1$  として6) へ。ただし、この時  $p > P_2 + Th$  であれば9) へ。

9) 終了。

25) このようにして6) ~ 8) の処理を  $P_2 - Th \sim P_2 + Th$  まで行って、相関の最も大きいもの  $C_{max}$  と仮ピッチ  $Q_1$ ,  $Q_2$  を求める。この時の  $Q_1$ ,  $Q_2$  が第1サブフレームと第2サブフレームの仮ピッチである。

上記アルゴリズムにより、2つのサブフレームの相関を同時に評価しながら大きさに比較的差のない (差の最大は  $Th$  である) 仮ピッチを2つ選択するこ

- とができる。この仮ピッチを用いることにより、第2サブフレームの適応符号帳探索の際に、探索の範囲を狭く設定しても符号化性能を大きく劣化させることを防止できる。例えば、第2サブフレームから音質が急に変化した場合などで、第2サブフレームの相関が強い場合は、第2サブフレームの相関を反映した $Q_1$ を用いることで第2サブフレームの劣化を回避出来る。

さらに、探索範囲設定部311は、求めた仮ピッチ $Q_1$ を用いて適応符号帳の探索を行う範囲( $L_{ST} \sim L_{EN}$ )を下記式8のようにして設定する。

第1サブフレーム

- $$L_{ST} = Q_1 - 5 \quad (\text{ただし、} L_{ST} < L_{min} \text{の時 } L_{ST} = L_{min})$$
- $$L_{EN} = L_{ST} + 20 \quad (\text{ただし、} L_{ST} > L_{max} \text{の時 } L_{ST} = L_{max})$$

第2サブフレーム

- $$L_{ST} = T_1 - 10 \quad (\text{ただし、} L_{ST} < L_{min} \text{の時 } L_{ST} = L_{min})$$
- $$L_{EN} = L_{ST} + 21 \quad (\text{ただし、} L_{ST} > L_{max} \text{の時 } L_{ST} = L_{max})$$

式8

- 15 ここで、
- $L_{ST}$  : 探索範囲の最小
- $L_{EN}$  : 探索範囲の最大
- $L_{min}$  : ラグの最小値 (例 : 20)
- $L_{max}$  : ラグの最大値 (例 : 143)
- 20  $T_1$  : 第1フレームの適応符号帳ラグ

上記設定において、第1サブフレームは探索範囲を狭める必要はない。しかしながら、本発明者らは、入力音声のピッチに基づいた値の付近を探索区間とした方が性能が良いことを実験により確認しており、本実施の形態では26サンプルに狭めて探索するアルゴリズムを使用している。

- 25 また、第2サブフレームは第1サブフレームで求められたラグ $T_1$ を中心にその付近に探索範囲を設定している。したがって、合計32エントリで、第2サブフレームの適応符号帳のラグを5ビットで符号化できることになる。また、

本発明者らは、この時もラグの小さい候補を少なく、ラグの大きい候補を多く設定することにより、より良い性能が得られることを実験により確認している。ただし、これまでの説明でわかるように、本実施の形態においては、仮ピッチ  $Q_2$  は使用しない。

5       ここで、本実施の形態における効果について説明する。探索範囲設定部 3 1 1 によって得られた第 1 サブフレームの仮ピッチの近くには、第 2 サブフレームの仮ピッチも存在している（定数  $T_h$  で制限したため）。また、第 1 サブフレームにおいて探索範囲を絞って探索しているので、探索の結果得られるラグは第 1 サブフレームの仮ピッチから離れない。

10       したがって、第 2 サブフレームの探索の時には、第 2 サブフレームの仮ピッチから近い範囲を探索できることになり、第 1、第 2 サブフレームの両方において適当なラグが探索できることになる。

例として、第 1 サブフレームが無音で、第 2 サブフレームから音声が発生した場合を考える。従来法では、探索範囲を狭めることで第 2 サブフレームのピッチが探索区間に含まれなくなると、音質は大きく劣化してしまう。本実施の形態に係る方法においては、ピッチ分析部の仮ピッチの分析において、代表ピッチ  $P_2$  の相関は強く出る。したがって、第 1 サブフレームの仮ピッチは  $P_2$  付近の値になる。このため、デルタラグによる探索の際に、音声が発生した部分に近い部分を仮ピッチとすることができる。すなわち、第 2 サブフレームの適応符号帳の探索の時には、 $P_2$  付近の値を探索できることになり、途中で音声の立ち上がり生じて劣化なくデルダラグにより第 2 サブフレームの適応符号帳探索を行うことができる。

次に、音源作成部 3 0 5 において、適応符号帳 3 0 3 に格納された音源サンプル（適応コードベクトル又は適応音源）と確率的符号帳 3 0 4 に格納された音源サンプル（確率的コードベクトル又は確率的音源）を取り出し、それぞれを聴感重み L P C 合成部 3 0 6 へ送る。さらに、聴感重み L P C 合成部 3 0 6 において、音源作成部 3 0 5 で得られた 2 つの音源に対して、L P C 分析部 3

02で得られた復号化LPC係数によってフィルタリングを行って2つの合成音を得る。

さらに、ゲイン演算部308においては、聴感重みLPC合成部306で得られた2つの合成音と入力音声との関係を分析し、2つの合成音の最適値（最適ゲイン）を求める。また、ゲイン演算部308においては、その最適ゲインによってパワ調整したそれぞれの合成音を加算して総合合成音を得る。そして、ゲイン演算部308は、その総合合成音と入力音声の符号化歪みの計算を行う。また、ゲイン演算部308においては、適応符号帳303と確率的符号帳304の全ての音源サンプルに対して音源作成部305、聴感重みLPC合成部306を機能させることによって得られる多くの合成音と入力音声との間の符号化歪みを行い、その結果得られる符号化歪みの中で最も小さいときの音源サンプルのインデックスを求める。

次に、得られた音源サンプルのインデックス、そのインデックスに対応する2つの音源、及び入力音声をパラメータ符号化部309へ送る。パラメータ符号化部309では、ゲインの符号化を行うことによってゲイン符号を得て、LPC符号、音源サンプルのインデックスと共に伝送路へ送る。

また、パラメータ符号化部309は、ゲイン符号と音源サンプルのインデックスに対応する2つの音源から実際の音源信号を作成し、それを適応符号帳303に格納すると同時に古い音源サンプルを破棄する。

20    なお、聴感重みLPC合成部306においては、LPC係数や高域強調フィルタや長期予測係数（入力音声の長期予測分析を行うことによって得られる）を用いた聴感重み付けフィルタを用いる。

上記ゲイン演算部308は、音源作成部305から得られた適応符号帳303、確率的符号帳304の全ての音源について入力音声との間の比較を行うが、25    計算量削減のため、2つの音源（適応符号帳303と確率的符号帳304）については上述したようにしてオープンループにより探索する。

このように、本実施の形態におけるピッチ探索方法により、最初のサブフレ

ームの適応符号帳探索の前に、フレームを構成する複数のサブフレームのピッチ分析を行って相関値を算出することにより、フレーム内の全サブフレームの相関値を同時に把握することができる。

- そして、各サブフレームの相関値を算出すると共に、その相関値の大小から
- 5 各サブフレームで最もピッチ周期らしい値（代表ピッチと呼ぶ）を求め、ピッチ分析で得られた相関値と代表ピッチに基づいて複数のサブフレームのラグの探索範囲を設定する。この探索範囲の設定においては、ピッチ分析で得た複数のサブフレームの代表ピッチと相関値を利用して探索範囲の中心となる差の少ない適当な仮のピッチ（仮ピッチと呼ぶ）を求める。
- 10 さらに、上記探索範囲の設定で求めた仮ピッチの前後の指定の範囲にラグの探索区間を限定するので、適応符号帳の効率の良い探索を可能にする。その際、ラグの短い部分の候補を少なくし、ラグのより長い範囲を広く設定するので、良好な性能が得られる適当な探索範囲を設定することができる。また、適応符号帳探索の際に上記探索範囲の設定で設定された範囲でラグの探索を行うの
- 15 で、良好な復号化音を得ることができる符号化が可能になる。

- このように、本実施の形態によれば、探索範囲設定部 311 によって得られた第1サブフレームの仮ピッチの近くには第2サブフレームの仮ピッチも存在しており、第1サブフレームにおいて探索範囲を絞っているので、探索の結果得られるラグは仮ピッチから離れて行かない。したがって、第2サブフレームの探索の時には第2サブフレームの仮ピッチ付近を探索できることになり、
- 20 フレームの後半から音声が始まる場合などの非定常なフレームでも、第1、第2サブフレームにおいて適当なラグ探索が可能になり、従来得られなかった格別の効果を得ることができる。

（実施の形態3）

- 25 初期のCELP方式では、ランダム数列が確率的音源ベクトルとして複数種類エンタリーされた確率的符号帳、すなわち複数種類のランダム数列をメモリに直接記録した確率的符号帳が使われていた。一方、近年の低ビットレートC



ELP符号化・復号化装置においては、振幅が+1か-1の非零要素（非零要素以外の要素の振幅は零）を少数個含んだ確率的音源ベクトルを生成する代数的符号帳を確率的符号帳部に備えるものが多く開発されている。

- なお、代数的符号帳は、「Fast CELP Coding based on Algebraic codes」,  
5 J.Adoul et al, Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing, 1987, pp. 1957-1960や「Comparison of Some Algebraic Structure for CELP Coding of Speech」, J.Adoul et al, Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing, 1987, pp. 1953-1956などに開示されている。
- 10 上記文献に開示されている代数的符号帳は、（１）ビットレートが8 kb/s程度のCELP方式に適用した場合、品質の高い合成音を生成できる、（２）少ない演算量で確率的音源符号帳を探索できる、（３）確率的音源ベクトルを、直接格納しておくデータROM容量が不要になる、といった優れた特徴を有する符号帳である。
- 15 そして、代数符号帳を確率的符号帳として用いることを特徴とするCS-ACELP（ビットレート8 kb/s）やACELP（ビットレート5.3 kb/s）が、G. 729、g 723.1として、それぞれITU-Tから1996年に勧告化されている。なお、CS-ACELPに関しては、「Design and  
Description of CS-ACELP: A Toll Quality 8 kb/s Speech Coder」,  
20 Redwan Salami et al, IEEE trans. SPEECH AND AUDIO PROCESSING, vol. 6, no. 2, March 1998などに、その詳細技術が開示されている。

代数的符号帳は、上記のように優れた特徴を有する符号帳である。しかしながら、代数的符号帳をCELP符号化・復号化装置の確率的符号帳に適用した場合、確率的音源ターゲットは、比零要素を少数個だけ含んだ確率的音源ベクトルで常に符号化（ベクトル量子化）されることになるので、確率的音源ターゲットの忠実な符号表現は不可能であるという課題も生じている。そして、処理フレームが、無声子音区間や背景雑音区間などに相当する場合に、この課題

25

は特に顕著になる。

無声子音区間や背景雑音区間では、確率的音源ターゲットが複雑な形状になることが多いためである。またさらには、ビットレートが8 kb/s 程度よりさらに低いCELP符号化・復号化装置に代数的符号帳を適用した場合には、確

5 率的音源ベクトル中の比零要素数を少なくすることになるため、確率的音源ターゲットがパルスの形状になりやすい有声区間でさえも、上記課題が問題になる場合がある。

代数的符号帳の有する上記課題を解決する一方法として、代数的符号帳より出力される少数個の非零要素（非零要素以外の要素はゼロの値を持つ）を含む

10 ベクトルと、拡散パターンと呼ばれる固定波形とを重畳して得られるベクトルを、合成フィルタの駆動音源とするパルス拡散符号帳を用いる方法が開示されている。パルス拡散符号帳は、特開平10-232696号公報、「パルス拡散構造音源を併用するACELP符号化」安永他，電子情報通信学会平成9年度春季全国大会発表予稿集，D-14-11，p. 253，1997-03、「パルス拡散音

15 源を用いた低レート音声符号化」安永他，日本音響学会平成10年秋期研究発表会講演論文集，pp. 281-282，1998-10などに開示されている。

そこで次に、上記文献で開示されたパルス拡散符号帳の概要を、図8及び図9を用いて説明する。なお、図9は、図8のパルス拡散符号帳のさらに詳細な一例を示すものである。

20 図8及び図9のパルス拡散符号帳において、代数的符号帳4011は、少数個の非零要素（振幅は+1又は-1）からなるパルスベクトルを生成する符号帳である。上記文献に記載されているCELP符号化装置・復号化装置では、代数的符号帳4011の出力であるパルスベクトル（少数個の非零要素によって構成される）がそのまま、確率的音源ベクトルとして用いられている。

25 拡散パターン格納部4012には、拡散パターンと呼ばれる固定波形が、各チャネルあたり1種類以上ずつ格納されている。なお、各チャネルごとに格納された前記拡散パターンは、チャネル毎で異なる形状の拡散パターンが格納される場合、

各チャンネルに同一形状（共通の）の拡散パターンが格納される場合の双方が考えられる。各チャンネル用に格納される拡散パターンが共通の場合は、各チャンネル用に格納される拡散パターンが格納される場合を単純化したものに相当するので、本明細書の以下の説明では、チャンネル毎に格納される拡散パタンの形状がそれ

5 ぞれ異なる場合について説明を進めることとする。

パルス拡散符号帳 401 は、代数的符号帳 4011 からの出力ベクトルをそのまま確率的音源ベクトルとして出力するのではなく、代数的符号帳 4011 から出力されるベクトルと、拡散パターン格納部 4012 から読み出される拡散パターンとを、パルス拡散部 4013 でチャンネルごとに重畳し、重畳演算によって得られるベクトルを加算して得られるベクトルを確率的音源ベクトルとして利用する。

なお、上記文献において開示されている CELP 符号化・復号化装置は、符号化装置と復号化装置で同一構成（代数的符号帳部のチャンネル数、拡散パターン格納部に登録されている拡散パタンの種類数および形状などが、符号化装置側

15 と復号化装置側で共通）のパルス拡散符号帳を用いることを特徴としている。そして、拡散パターン格納部 4012 に登録しておく拡散パタンの形状、種類数、複数種類以上登録している場合にはそれらの選択方法を効率的に設定することによって、合成音声の品質を向上を図っている。

なお、パルス拡散符号帳に関するここでの説明は、少数個の非零要素からなるパルスベクトルを生成する符号帳として、非零要素の振幅を +1 もしくは -1 に限定した代数的符号帳を用いた場合についての説明であるが、当該パルスベクトルを生成する符号帳としては、非零要素の振幅を限定しないマルチパルス符号帳や、レギュラーパルス符号帳を用いることも可能であり、その場合にも、パルスベクトルを拡散パターンと重畳したものを確率的音源ベクトルとして

25 利用することで合成音声の品質向上を実現できる。

これまでに、多くの確率的音源ターゲットの形状を統計学習し、確率的音源ターゲット中に統計的に高い頻度で含まれる形状の拡散パターン、無声子音区間

や雑音区間を効率的に表現するための乱数的な形状の拡散パターン、有声定常区間を効率的に表現するためのパルス的な形状の拡散パターン、代数的符号帳から出力されるパルスベクトルのエネルギー（非零要素の位置にエネルギーが集中

- している）を周囲に分散させるような作用を与える形状の拡散パターン、適当に
- 5 用意したいくつかの拡散パターン候補について、音声信号を、符号化、復号化、合成音声の視聴評価を繰り返し、品質の高い合成音声を出力しうるよう選択した拡散パターン、又は音声学的な知見をもとに作成した拡散パターンなどを、代数的符号帳から出力される音源ベクトル中の非零要素（チャンネル）あたり1種類以上ずつ登録しておき、登録しておいた拡散パターンと、代数的符号帳によって
- 10 生成されるベクトル（少数個の非零要素によって構成される）とをチャンネルごとに重畳し、各チャンネルの重畳結果を加算したものを確率的音源ベクトルとして用いることにより、合成音声を有効に品質向上させることができることが開示されてきた。

- また、特に、拡散パターン格納部4012が、チャンネルあたり複数種類（2種類以上）の拡散パターンを登録している場合については、それら複数の拡散パ
- 15 ターの選択方法として、登録された拡散パターの全組合わせについて実際に符号化・復号化を行い、その結果生じる符号化歪みが最小になるような拡散パターンをクローズド選択する方法や、確率的符号帳探索を行う時点で既に明らかになっている音声的情報（ここでいう音声的情報とは、例えば、ゲイン符号の動的
- 20 変動情報もしくはゲイン値の（予め設定したしきい値との）大小関係情報などを利用して判定した有声性の強弱情報、あるいは、線形予測符号の動的変動を利用して判定した有声性の強弱情報などのことである）を利用して、拡散パターンをオープン選択する方法などが開示されている。

- なお、以降の説明では、説明を簡単にするため、図9のパルス拡散符号帳内
- 25 の拡散パターン格納部4013が、チャンネルあたり1種類だけの拡散パターンを登録していることを特徴とする図10のパルス拡散符号帳に限定して説明する。

ここでは次に、代数的符号帳をCELP符号化装置に適用した場合の確率的

符号帳探索処理と比較して、パルス拡散符号帳をCELP符号化装置に適用した場合の確率的符号帳探索処理を説明する。まず、代数的符号帳を確率的符号帳部に用いた場合の符号帳探索処理を説明する。

- 代数的符号帳によって出力されるベクトル内の非零要素数をN（代数的符号帳のチャンネル数をN）、チャンネルごとに出力する振幅が+1か-1の非零要素を1本だけ含むベクトル（非零要素以外の要素の振幅はゼロ）を $d_i$ （ $i$ はチャンネル番号： $0 \leq i \leq N-1$ ）、サブフレーム長をLとした時、代数的符号帳によって出力されるエントリ番号 $k$ の確率的音源ベクトル $c_k$ は、下記式9となる。

10

$$C_k = \sum_{i=0}^{N-1} d_i$$

$C_k$ : 代数的符号帳によるエントリ番号  $K$  の確率的音源ベクトル

$d_i$ : 非零要素ベクトル ( $d_i = \pm \delta(n - p_i)$ ) ただし、 $p_i$ : 非零要素位置

$N$ : 代数的符号帳のチャンネル数 (= 確率的音源ベクトル中の非零要素数)

15

式9

そして、式9を式10に代入することで、下記式11が得られる。

$$D_k = \frac{(v^t H c_k)^2}{\|H c_k\|^2}$$

$v^t$  :  $v$  (確率的音源ターゲット) の転置ベクトル

20  $H^t$  :  $H$  (合成フィルタのインパルス応答行列) の転置行列

$c_k$  : エントリ番号  $k$  番目の確率的音源ベクトル

式10

$$D_k = \frac{\left( v^t H \left( \sum_{i=0}^{N-1} d_i \right) \right)^2}{\left\| H \left( \sum_{i=0}^{N-1} d_i \right) \right\|^2}$$

$v$ : 確率的音源ターゲットベクトル

$H$ : 合成フィルタのインパルス応答畳み込み行列

$d_i$ : 非零要素ベクトル ( $d_i = \pm \delta(n - p_i)$ ) ただし、 $p_i$ : 非零要素位置

5  $N$ : 代数的符号帳のチャネル数 (= 確率的音源ベクトル中の非零要素数)

$$x^t = v^t H$$

$$M = H^t H$$

式 1 1

この式 1 0 を整理して得られる下記式 1 2 を最大化するようなエントリ番号  $k$  を特定する処理が確率的符号帳探索処理となる。

$$D_k = \frac{\left( \sum_{i=0}^{N-1} x^t d_i \right)^2}{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} d_i^t M d_j}$$

式 1 2

ただし、式 1 2 において、 $x^t = v^t H$ 、 $M = H^t H$  ( $v$  は確率的音源ターゲット) である。ここで各エントリ番号  $k$  について式 1 2 の値を計算する場合、その前  
15 処理段階で  $x^t = v^t H$  及び  $M = H^t H$  を計算し、計算結果をメモリに展開 (記憶) させておく。この前処理を導入することで、確率的音源ベクトルとしてエントリしている各候補ごとに式 1 2 を計算する際の演算量を大幅に削減でき、この結果として、確率的符号帳探索に要するトータルの演算量を少なくおさえられることが、上記文献などに開示されており、一般に知られている。

20 次に、パルス拡散符号帳を確率的符号帳に用いた場合の確率的符号帳探索処理を説明する。

- パルス拡散符号帳の構成一部位である代数的符号帳によって出力される非  
 零要素数を $N$ （代数的符号帳のチャンネル数を $N$ ）、チャンネルごとに出力する振  
 幅が $+1$ か $-1$ の非零要素を1本だけ含むベクトル（非零要素以外の要素の振  
 幅はゼロ）を $d_i$ （ $i$ はチャンネル番号： $0 \leq i \leq N-1$ ）、拡散パターン格納部  
 5 が格納しているチャンネル番号 $i$ 用の拡散パターンを $w_i$ 、サブフレーム長を $L$ と  
 した時、パルス拡散符号帳によって出力されるエン트리番号 $k$ の確率的音源  
 ベクトル $c_k$ は、下記式13となる。

$$C_k = \sum_{i=0}^{N-1} W_i d_i$$

- 10  $C_k$ : パルス拡散符号帳によるエン트리番号 $K$ の確率的音源ベクトル

$W_i$ : 拡散パターン( $w_i$ )重畳行列

$d_i$ : 代数的符号帳部が出力する非零要素ベクトル

( $d_i = \pm \delta(n - p_i)$  ただし、 $p_i$ : 非零要素位置)

$N$ : 代数的符号帳部のチャンネル数

15

式13

したがってこの場合、式13を式10に代入することで、下記式14が得ら  
 れる。

$$D_k = \frac{\left( v^H \left( \sum_{i=0}^{N-1} W_i d_i \right) \right)^2}{\left\| H \left( \sum_{i=0}^{N-1} W_i d_i \right) \right\|^2}$$

$v$ : 確率的音源ターゲットベクトル

- 20  $H$ : 合成フィルタのインパルス応答畳み込み行列

$W_i$ : 拡散パターン( $w_i$ )重畳行列

$d_i$ : 代表的符号帳部が出力する非零要素ベクトル

( $d_i = \pm \delta(n - p_i)$  ただし、 $p_i$ : 非零要素位置)

$N$ : 代数的符号帳のチャンネル数 (= 確率的音源ベクトル中の非零要素数)

$$H_i = HW_i$$

$$x_i^t = v^t H_i$$

$$R = H_i H_j$$

式 1 4

- 5 この式 1 4 を整理して得られる下記式 1 5 を最大化する確率的音源ベクトルのエントリ番号  $k$  を特定する処理が、パルス拡散符号帳を用いた場合の確率的符号帳探索処理となる。

$$D_k = \frac{\left( \sum_{i=0}^{N-1} x_i^t d_i \right)^2}{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} d_i^t R d_j}$$

式 1 5

- 10 ただし、式 1 5 において、 $x^t = v^t H_i$  (ただし、 $H_i = HW_i$  :  $W_i$  は拡散パターン重畳行列)、である。各エントリ番号  $k$  について式 1 5 の値計算する場合、その前処理として  $H_i = HW_i$  及び  $x^t = v^t H_i$  及び  $R = H_i^t H_j$  を計算しメモリに記録しておくことが可能である。すると、確率的音源ベクトルとしてエントリしている各候補ごとに式 1 5 を計算する際の演算量が、代数的符号帳を用いた場合に
- 15 式 1 2 を計算する際の演算量と同じになり (式 1 2 と式 1 5 が同形であることから明らか)、パルス拡散符号帳を用いた場合も、少ない演算量で確率的符号帳探索を行うことができる。

- 上記技術においては、パルス拡散符号帳を CELP 符号化装置・復号化装置の確率的符号帳部に用いることの効果、及びパルス拡散符号帳を確率的符号帳部に用いた場合に、代数的符号帳を確率的符号帳部に用いた場合と同様の方法で確率的符号帳探索を行えることを示した。代数的符号帳を確率的符号帳部に用いた場合の確率的符号帳探索に要する演算量と、パルス拡散符号帳を確率的符号帳部に用いた場合の確率的符号帳探索に要する演算量の違いは、式 1 2 と
- 20 式 1 5 それぞれの前処理段階に要する演算量の違い、すなわち、前処理



( $\mathbf{x}^t = \mathbf{V}^t \mathbf{H} \mathbf{i}$ ,  $\mathbf{M} = \mathbf{H}^t \mathbf{H}$ ) と前処理 ( $\mathbf{H} \mathbf{i} = \mathbf{H} \mathbf{W} \mathbf{i}$ ,  $\mathbf{x}^t = \mathbf{V}^t \mathbf{H} \mathbf{i}$ ,  $\mathbf{R} = \mathbf{H} \mathbf{i}^t \mathbf{H} \mathbf{j}$ ) に要する演算量の違いである。

- 一般に、CELP符号化装置・復号化装置では、そのビットレートが低くなるほど確率的符号帳部に割り当て可能なビット数も減少する傾向にある。そしてこの傾向は、代数的符号帳やパルス拡散符号帳を確率的符号帳部に用いる場合、確率的音源ベクトルを構成する際の非零要素数の減少につながっていく。したがって、CELP符号化装置・復号化装置のビットレートが低くなるほど、代数的符号帳を用いた場合とパルス拡散符号帳を用いた場合の演算量の差は少なくなる。しかしビットレートが比較的高い場合や、ビットレートが低くても演算量を極力少なく押さえる必要がある場合には、パルス拡散符号帳を用い
- 5
- 10
- ることによって生じる前処理段階の演算量の増加が無視できなくなることがある。

- 本実施の形態では、パルス拡散符号帳を確率的符号帳部に用いたCELP方式の音声符号化装置と音声復号化装置、及び音声符号化復号化システムにおいて、代数的符号帳を確率的符号帳部に用いる場合と比べて増加する、符号探索処理における前処理部分の演算量増加分を少なく抑えながら、復号化側では高品質な合成音声を得ることについて説明する。
- 15

- 具体的には、本実施の形態に係る技術は、パルス拡散符号帳をCELP符号化装置・復号化装置の確率的符号帳部に用いる場合に生じることがある上記課題を解決するためのものであり、符号化装置側と復号化装置側で異なる拡散パターンを用いることを特徴である。すなわち、本実施の形態においては、音声復号化装置側の拡散パターン格納部には、上述した拡散パターンを登録し、それを用いることで、代数的符号帳を用いる場合より品質の高い合成音声を生成する。
- 20
- 一方、音声符号化装置側では、復号化装置側の拡散パターン格納部に登録する拡散パターンを簡素化した拡散パターン（例えば、一定間隔で間引いた拡散パターンや、ある長さで打ち切った拡散パターン）を登録し、それを用いて確率的符号帳探索を行うようにする。
- 25

これにより、パルス拡散符号帳を確率的符号帳部に用いる場合に、符号化側では、代数的符号帳を確率的符号帳部に用いる場合と比べて増加する、前処理段階の符号探索時の演算量を少なく抑えることができ、~~復号化側では、高品質~~の合成音声を得ることができる。

- 5      符号化装置側と復号化装置側で異なる拡散パターンを用いることは、予め用意された（復号化装置用の）拡散ベクトルを、その特性を残しつつ変形することにより、エンコーダ用の拡散ベクトルを獲得することである。

- ここで、復号化装置用の拡散ベクトルを予め用意する方法としては、本発明者らが以前に出願した特許（特開平10-63300号公報）に開示された方法、すなわち音源探索用ターゲットベクトルの統計的傾向を学習することによって用意する方法、音源ターゲットを実際に符号化し、その時生じる符号化歪みの総和をより小さくする方向に徐々に変形させる操作を反復することで用意する方法、及び合成音声を高品質化すべく音声学的な知見に基づいて設計する方法などや、パルス音源の高域位相成分をランダム化させることを目的に
- 10      設計する方法などが考えられる。これらの内容はすべてここに含めておく。

- このようにして得られた拡散ベクトルは、いずれも拡散ベクトルの先頭サンプルに近いサンプル（前方のサンプル）の振幅が、後方のサンプルの振幅より、比較的大きめになるという特徴がある。中でも、先頭のサンプルの振幅が、拡散ベクトル内の全サンプル中で最大となることが多い（ほとんどの場合そのようになる）。
- 20

復号化装置用の拡散ベクトルを、その特性を残しつつ変形することでエンコーダ用の拡散ベクトルを獲得する具体的方法としては、以下の方法が挙げられる。

- 1) 復号化装置用の拡散ベクトルのサンプル値を、適当な間隔ごとにゼロに置き換えることで、エンコーダ用の拡散ベクトルを獲得する。
- 25

2) ある長さの復号化装置用の拡散ベクトルを、適当な長さで打ち切ることによって、エンコーダ用の拡散ベクトルを獲得する。

3) 振幅のしきい値を予め設定し、復号化装置用の拡散ベクトルに対して設定したしきい値より振幅の小さいサンプルをゼロに置き換えることで、エンコーダ用の拡散ベクトルを獲得する。

4) ある長さの復号化装置用の拡散ベクトルを、先頭サンプルを含む適当な  
5 間隔ごとのサンプル値を保存し、それ以外のサンプルの値をゼロに置きかえることで、符号化装置用の拡散ベクトルを獲得する。

ここで例えば上記1)の方法のように、拡散ベクトルの前方からの数サンプルを用いた場合でも、拡散ベクトルの概形（大まかな特性）を保存したまま、符号化装置用の拡散ベクトルを新たに獲得することが可能となっている。

10 また例えば、上記2)の方法のように、適当な間隔ごとにサンプル値をゼロに置き換えてももとの拡散ベクトルの概形（大まかな特性）を保存したまま、符号化装置用の拡散ベクトルを新たに獲得することが可能となる。特に、上記  
4)の方法の場合は、振幅が最大であることの多い先頭サンプルの振幅をそのまま必ず保存するという限定を付けているので、もとの拡散ベクトルの概形を  
15 より確実に保存しておくことが可能である。

また、3)の方法のように、特定値以上の振幅を有するサンプルをそのまま保存し、前記特定値以下の振幅を有するサンプルの振幅をゼロに置き換えても拡散ベクトルの概形（大まかな特性）を保存したまま、符号化装置用の拡散ベクトルを獲得することが可能となる。

20 以下、本実施の形態に係る音声符号化装置及び音声復号化装置について、添付図面を参照して詳細に説明する。なお、添付図面に記載のCELP音声符号化装置（図11）、および、CELP音声復号化装置（図12）は、従来のCELP音声装置およびCELP音声復号化装置における確率的符号帳部分に、上記のパルス拡散符号帳を用いている点に特徴を有している。従って、以降の  
25 説明において、確率的符号帳、確率的音源ベクトル、確率的音源ゲインと記載された部分は、それぞれ、パルス拡散符号帳、パルス拡散音源ベクトル、パルス拡散音源ゲインと読み替えることが可能である。なお、CELP音声符号化

装置およびCELP音声復号化装置における確率的符号帳は、雑音符号帳、あるいは、複数種類の固定波形を格納する作用を有することから固定符号帳と呼ばれることもある。

- 5        図11のCELP音声符号化装置では、まず始めに、線形予測分析部501が、入力音声線を線形予測分析して線形予測係数を算出し、算出した線形予測係数を線形予測係数符号化部502へ出力する。次に、線形予測係数符号化部502が、線形予測係数を符号化（ベクトル量子化）し、ベクトル量子化によって得られる量子化インデックス（以下、線形予測符号と呼ぶ）を符号出力部513及び線形予測符号復号化部503へ出力する。

次いで、線形予測符号復号化部503が、線形予測係数符号化部502で得た線形予測符号を復号化（逆量子化）して合成フィルタ504へ出力する。合成フィルタ504は、線形予測符号復号化部503で復号化して得られた復号化線形予測符号を係数に持つ全極型モデルの合成フィルタを構成する。

- 15        そして、適応符号帳506から選出される適応音源ベクトルに適応音源ゲイン509を乗じて得られるベクトルと、パルス拡散符号帳507から選出した確率的音源ベクトルに確率的音源ゲイン510を乗じて得られるベクトルとをベクトル加算部511で加算して駆動音源ベクトルを生成する。そして、歪み計算部505が、当該駆動音源ベクトルで合成フィルタ504を駆動したときの出力ベクトルと、入力音声との歪みを下記式16により計算し、歪みERを符号特定部512へ出力する。

$$ER = \|u - (g_a H p + g_c H c)\|^2$$

u: 入力音声 (ベクトル)

H: 合成フィルタのインパルス応答行列

- 25    p: 適応音源ベクトル

c: 確率的音源ベクトル

$g_a$ : 適応音源ゲイン

$g_c$ : 確率的音源ゲイン

式 1 6

ただし、式 1 6 において、 $u$  は処理フレーム内の入力音声ベクトル、 $H$  は合成フィルタのインパルス応答行列、 $g_a$  は適応音源ゲイン、 $g_c$  は確率的音源ゲイン、 $p$  は適応音源ベクトル、 $c$  は確率的音源ベクトルを示す。

ここで、適応符号帳 5 0 6 は、過去数フレーム分の駆動音源ベクトルを格納したバッファ（動的メモリ）であり、上記適応符号帳 5 0 6 から選出される適応音源ベクトルは、入力音声を合成フィルタの逆フィルタに通して得られる線形予測残差ベクトル中の周期成分を表現するために使われる。

一方、パルス拡散符号帳 5 0 7 から選出される音源ベクトルは、線形予測残差ベクトルに現処理フレームで新たに加わった非周期成分（線形予測残差ベクトルから周期性（適応音源ベクトル成分）を除去した成分）を表現するために使われる。

そして、適応音源ベクトルゲイン乗算部 5 0 9 及び確率的音源ベクトルゲイン乗算部 5 1 0 は、適応符号帳 5 0 6 から選出される適応音源ベクトル及びパルス拡散符号帳 5 0 7 から選出される確率的音源ベクトルに対して、ゲイン符号帳 5 0 8 から読みだした適応音源ゲイン及び確率的音源ゲインを乗じる機能を有している。なお、ゲイン符号帳 5 0 8 とは、適応音源ベクトルに乗じる適応音源ゲインと、確率的音源ベクトルに乗じる確率的音源ゲインとのセットを複数種類格納した静的メモリである。

符号特定部 5 1 2 は、歪み計算部 5 0 5 で計算した式 1 6 の歪み  $E_R$  を最小化する上記 3 つの符号帳（適応符号帳、パルス拡散符号帳、ゲイン符号帳）のインデックスの最適組み合わせを選択する。そして、歪み特定部 5 1 2 は、上記歪みが最小になるときに選択していた各符号帳のインデックスを、それぞれ適応音源符号、確率的音源符号、ゲイン符号として符号出力部 5 1 3 へ出力する。

そして最後に、符号出力部 5 1 3 は、線形予測係数符号化部 5 0 2 で得られ

た線形予測符号と、符号特定部 5 1 2 で特定された適応音源符号、確率的音源符号及びゲイン符号を、全てまとめて現処理フレーム内の入力音声表現する符号（ビット情報）とし、復号化装置側へ出力する。

- 5      なお、符号特定部 5 1 2 で行う適応音源符号、確率的音源符号、ゲイン符号の特定は、一定時間間隔のフレームを、サブフレームと呼ぶさらに短い時間間隔に分割した上で行われることがある。ただし、本明細書では、フレームとサブフレームと特に区別しないで（フレームという呼び方に統一した上で）、以下の説明を行う。

次に、CELP 音声復号化装置の概要を、図 1 2 を用いて説明する。

- 10      図 1 2 の CELP 復号化装置では、まず、符号入力部 6 0 1 が、CELP 音声符号化装置（図 1 1）で特定した符号（フレーム区間内の音声信号を符号表現するためのビット情報）を受け、受けた符号を線形予測符号、適応音源符号、確率的音源符号、及びゲイン符号の 4 種類の符号に分解する。そして、線形予測符号を線形予測係数復号化部 6 0 2 へ、適応音源符号を適応符号帳 6 0 3 へ、  
15      確率的音源符号をパルス拡散符号帳 6 0 4 へ、ゲイン符号をゲイン符号帳 6 0 5 へ出力する。

次に、線形予測係数復号化部 6 0 2 は、符号入力部 6 0 1 から入力される線形予測符号を復号化して復号化線形予測符号を得て、この復号化線形予測符号を合成フィルタ 6 0 9 へ出力する。

- 20      合成フィルタ 6 0 9 は、線形予測係数復号化部 6 0 2 で得た復号化線形予測符号を係数にもつ全極型モデルの合成フィルタを構成する。また、適応符号帳 6 0 3 は、符号入力部 6 0 1 から入力された適応音源符号に対応する適応音源ベクトルを出力する。また、パルス拡散符号帳 6 0 4 は、符号入力部 6 0 1 から入力された確率的音源符号に対応する確率的音源ベクトルを出力する。また、  
25      ゲイン符号帳 6 0 5 は、符号入力部 6 0 1 から入力されるゲイン符号に対応する適応音源ゲイン及び確率的音源ゲインを読み出し、それぞれ適応音源ゲイン乗算部 6 0 6 及び確率的音源ゲイン乗算部 6 0 7 へ出力する。

そして、適応音源ゲイン乗算部 606 が、適応符号帳 603 から出力された適応音源ベクトルに、ゲイン符号帳 605 から出力された適応音源ゲインを乗算し、確率的音源ゲイン乗算部 607 が、パルス拡散符号帳 604 から出力された確率的音源ベクトルに、ゲイン符号帳 605 で出力された確率的音源ゲインを乗算する。そしてベクトル加算部 608 が、適応音源ゲイン乗算部 606 及び確率的音源ゲイン乗算部 607 それぞれの出力ベクトルを加算して駆動音源ベクトルを生成する。そして、当該駆動音源ベクトルで、合成フィルタ 609 を駆動し、受信したフレーム区間の合成音声出力する。

このような CELP 方式の音声符号化装置・音声復号化装置において、品質の高い合成音声を得るためには、式 16 の歪み  $ER$  を小さく抑えることが必要になる。そのためには、式 16 の  $ER$  を最小化するように、適応音源符号、確率的音源符号、ゲイン符号の組み合わせを閉ループで特定することが望ましい。しかしながら、式 16 の歪み  $ER$  をクローズドループで特定しようとする演算処理量が大きくなりすぎるため、上記 3 種類の符号はオープンループで特定していくことが一般的である。

具体的には、まず、適応符号帳探索を行う。ここで、適応符号帳探索処理とは、入力音声を逆フィルタに通して得られる予測残差ベクトル中の周期性成分を、過去フレームの駆動音源ベクトルを格納した適応符号帳から出力される適応音源ベクトルによってベクトル量子化する処理である。そして、線形予測残差ベクトル内の周期成分と、近い周期成分を有する適応音源ベクトルのエンタリ番号を適応音源符号として特定する。なお、適応符号帳探索によって、同時に、理想適応音源ゲインが暫定的に確定されることになる。

次いで、パルス拡散符号帳探索を行う。パルス拡散符号帳探索は、処理フレームの線形予測残差ベクトルから周期成分を除去した成分、すなわち、線形予測残差ベクトルから適応音源ベクトル成分を差し引いた成分（以下、確率的音源ターゲットと呼ぶこともある）を、パルス拡散符号帳に格納された複数の確率的音源ベクトル候補を用いてベクトル量子化する処理である。そして、この

パルス拡散符号帳探索処理により、確率的音源ターゲットを、もっとも歪み少なく符号化する確率的音源ベクトルのエントリ番号を確率的音源符号として特定する。なお、パルス拡散符号帳探索によって、同時に、理想確率的ゲインも暫定的に確定されることになる。

- 5      そして最後に、ゲイン符号帳探索を行う。ゲイン符号帳探索は、適応符号帳探索時に暫定的に得られた理想適応ゲインと、パルス拡散符号帳探索時に暫定的に得られた理想確率的ゲインとの2要素からなるベクトルを、ゲイン符号帳に格納されたゲイン候補ベクトル（適応音源ゲイン候補と確率的音源ゲイン候補の2要素からなるベクトル候補）で歪みが最小になるように符号化（ベクトル量子化）する処理である。そして、ここで選択されるゲイン候補ベクトルの  
10      エントリ番号がゲイン符号として符号出力部へ出力される。

ここでは、次に、C E L P 音声符号化装置における上記一般的な符号探索処理のうち、パルス拡散符号帳探索処理（適応音源符号を特定した後に、確率的音源符号を特定する処理）についてさらに詳しく説明を行う。

- 15      説明したように、一般的なC E L P 符号化装置では、パルス拡散符号帳探索を行う時点では、線形予測符号及び適応音源符号は、既に特定されている。ここで、既に特定されている線形予測符号によって構成される合成フィルタのインパルス応答行列をH、適応音源符号と対応する適応音源ベクトルをp、適応音源符号を特定した時点で同時に求まる理想適応音源ゲイン（暫定値）をg<sub>a</sub>  
20      とすると、式16の歪みE<sub>R</sub>は、下記式17へと変形される。

$$ER_k = \|v - g_c H c_k\|^2$$

v: 確率的音源ターゲット(ただし、 $v = u - g_a H p$ )

$g_c$ : 確率的音源ゲイン

H: 合成フィルタのインパルス応答行列

- 25       $c_k$ : 確率的音源ベクトル(k: エントリー番号)



ただし、式 17 内のベクトル  $v$  は、フレーム区間内の入力音声信号  $u$ 、合成フィルタのインパルス応答行列  $H$ （既定）、適応音源ベクトル  $p$ （既定）、理想適応音源ゲイン  $g_a$ （暫定値）を用いた、下記式 18 の確率的音源ターゲットである。

$$5 \quad v = u - g_a H p$$

$u$ : 入力音声 (ベクトル)

$g_a$ : 適応音源ゲイン (暫定値)

$H$ : 合成フィルタのインパルス応答行列

$p$ : 適応音源ベクトル

10

式 18

なお、式 16 では確率的音源ベクトルが  $c$  と表現されており、一方、式 17 では確率的音源ベクトルは  $c_k$  と表現がされている。これは、式 16 では確率的音源ベクトルのエン트리番号 ( $k$  のこと) を違いを明示していないことに対して、式 17 ではエン트리番号を明示していることによるものであり、表現上の違いはあるものの意味する対象は同じものである。

したがって、パルス拡散符号帳探索とは、式 17 の歪み  $E R_k$  を最小化するような確率的音源ベクトル  $c_k$  のエン트리番号  $k$  を求める処理である。そして、式 17 の歪み  $E R_k$  を最小化するような確率的音源ベクトル  $c_k$  のエン트리番号  $k$  を特定する際には、確率的音源ゲイン  $g_c$  は任意の値をとりうると仮定できる。したがって、式 17 の歪みを最小化するようなエン트리番号を求める処理は、上記式 10 の分数式  $D_k$  を最大化するような確率的音源ベクトル  $c_k$  のエン트리番号  $k$  を特定する処理に置き換えられる。

25     そして、パルス拡散符号帳探索は、確率的音源ベクトル  $c_k$  のエン트리番号  $k$  ごとに式 10 の分数式  $D_k$  を歪み計算部 505 で計算し、その値を符号特定部 512 へ出力し、符号特定部 512 で、エン트리番号  $k$  ごとの式 10 の値を

大小比較して、その値が最大になるときのエントリ番号  $k$  を確率的音源符号と決定して符号出力部 513 へ出力する、といった 2 段階の処理によって行われることになる。

以下、本実施の形態における音声符号化装置及び音声復号化装置の動作について説明する。

図 11 に示す音声符号化装置におけるパルス拡散符号帳 507 の構成を図 13A に示し、図 12 に示す音声復号化装置におけるパルス拡散符号帳 604 の構成を図 13B に示す。図 13A に示すパルス拡散符号帳 507 と図 13B に示すパルス拡散符号帳 604 を比較した場合、構成上の異なる点は、拡散パターン格納部に登録している拡散パタンの形状が異なっている点である。

図 13B の音声復号化装置側では、拡散パターン格納部 4012 には、(1) 多くの確率的音源ターゲットの形状を統計学習し、確率的音源ターゲット中に統計的に高い頻度で含まれる形状の拡散パターン、(2) 無声子音区間や雑音区間を効率的に表現するための乱数的な形状の拡散パターン、(3) 有声定常区間を効率的に表現するためのパルスの形状の拡散パターン、(4) 代数的符号帳から出力される音源ベクトルのエネルギー（非零要素の位置にエネルギーが集中している）を周囲に分散させるような作用を与える形状の拡散パターン、(5) 適当に用意したいくつかの拡散パターン候補について、音声信号を、符号化、復号化、合成音声の視聴評価を繰り返し、品質の高い合成音声を出力しようよう選択した拡散パターン、(6) 音声学的な知見をもとに作成した拡散パタンのうちのいずれかの拡散パターンが各チャネルあたり 1 種類ずつ登録されている。

一方、図 13A の音声符号化装置側では、拡散パターン格納部 4012 には、図 13B の音声復号化装置側の拡散パターン格納部 4012 に登録されている拡散パターンを、1 サンプルおきにゼロに置き換えた拡散パターンが登録されている。

そして、上述のように構成された CELP 音声符号化装置／音声復号化装置では、符号化装置側と復号化装置側で異なる拡散パターンが登録されていること

を意識せずに、上記と同様の方法で、音声信号を符号化・復号化する。

符号化装置では、パルス拡散符号帳を確率的符号帳部に用いた場合の確率的符号帳探索時の前処理演算量を削減することができ ( $H_i = H_t W_i$  及び  $x_{it} = v_t H_i$  の演算量をおよそ半分に削減でき)、復号化装置側では、従来どおりの拡散パターンをパルスベクトルに重畳することで、非零要素位置に集中して  
5 いるエネルギーを周囲に拡散することができ、合成音声の品質を向上することが可能となる。

なお、本実施の形態では、図 13 A 及び図 13 B に示すように、音声符号化装置側では、音声復号化装置側で用いる拡散パターンを 1 サンプルおきにゼロに  
10 置き換えた拡散パターンを用いる場合について説明したが、音声符号化装置側では、音声復号化装置側で用いる拡散パターンの要素を  $N$  ( $N \geq 1$ ) サンプルおきにゼロに置き換えて得られる拡散パターンを用いた場合にも、本実施の形態をそのまま適用することができ、その場合にも同様の作用を得ることができる。

また、本実施の形態では、拡散パターン格納部が、チャンネルあたり 1 種類ずつ  
15 の拡散パターンを登録している場合の実施の形態を説明したが、チャンネルあたり 2 種類以上の拡散パターンが登録されており、それら拡散パターンを選択して用いることを特徴とするパルス拡散符号帳を確率的符号帳部に用いる CELP 音声符号化装置・復号化装置においても本発明を適用することが可能であり、その場合にも同様の作用・効果を得ることができる。

また、本実施の形態では、代数的符号帳部が 3 個の非零要素を含むベクトル  
20 を出力するパルス拡散符号帳を用いた場合について実施の形態を説明したが、代数的符号帳部が出力するベクトル中の非零要素数が  $M$  個 ( $M \geq 1$ ) の場合においても本実施の形態を適用することが可能であり、その場合にも同様の作用・効果を得ることができる。

また、本実施の形態では、少数個の非零要素からなるパルスベクトルを生成  
25 する符号帳として代数的符号帳を用いた場合について説明したが、当該パルスベクトルを生成する符号帳としては、マルチパルス符号帳やレギュラーパルス

符号帳など、その他の符号帳を用いる場合にも本実施の形態を適用することが可能であり、その場合にも同様の作用・効果を得ることができる。

次に、図 1 1 に示す音声符号化装置におけるパルス拡散符号帳の構成を図 1

4 A に示す、図 1 2 に示す音声復号化装置におけるパルス拡散符号帳の構成を

5 図 1 4 B に示す。

図 1 4 A に示すパルス拡散符号帳と図 1 4 B に示すパルス拡散符号帳の構成を比較した場合、構成上の異なる点は、拡散パターン格納部に登録している拡散パタンの長さが異なっている。図 1 4 B の音声復号化装置側では、拡散パターン格納部 4 0 1 2 には、上述した拡散パターンと同様の拡散パターン、すなわち、

- 10 (1) 多くの確率的音源ターゲットの形状を統計学習し、確率的音源ターゲット中に統計的に高い頻度で含まれる形状の拡散パターン、(2) 無声子音区間や雑音区間を効率的に表現するための乱数的な形状の拡散パターン、(3) 有声定常区間を効率的に表現するためのパルスの形状の拡散パターン、(4) 代数的符号帳から出力される音源ベクトルのエネルギー（非零要素の位置にエネルギーが集中している）を周囲に分散させるような作用を与える形状の拡散パターン、
- 15 (5) 適当に用意したいくつかの拡散パターン候補について、音声信号を、符号化、復号化、合成音声の視聴評価を繰り返し、品質の高い合成音声を出力しうよう選択した拡散パターン、(6) 音声学的な知見をもとに作成した拡散パタンのうちのいずれかの拡散パターンが各チャネルあたり 1 種類ずつ登録されて
- 20 いる。

一方、図 1 4 A の音声符号化装置側では、拡散パターン格納部 4 0 1 2 には、図 1 4 B の音声復号化装置側の拡散パターン格納部に登録されている拡散パターンを、半分の長さで打ち切った拡散パターンが登録されている。

- そして、上述のように構成された CELP 音声符号化装置・復号化装置では、
- 25 符号化装置側と復号化装置側で異なる拡散パターンが登録されていることを意識せずに、上述した場合と同様の方法で、音声信号を符号化・復号化する。

符号化装置では、パルス拡散符号帳を確率的符号帳部に用いた場合の確率的

符号帳探索時の前処理演算量を削減することができ ( $H_i = H_t W_i$  および  $x_{it} = v_t H_i$  の演算量をおよそ半分に削減でき)、復号化装置側では、従来どおりの拡散パターンを利用することで、合成音声の品質向上を実現することが可能となる。

- 5      なお、本実施の形態では、図 1 4 A 及び図 1 4 B に示すように、音声符号化装置側では、音声復号化装置側で用いる拡散パターンを半分の長さで打ち切った拡散パターンを用いる場合について説明したが、音声符号化装置側では、音声符号化装置側で用いる拡散パターンを、さらに短い長さ  $N$  ( $N \geq 1$ ) で打ち切った場合には、確率的符号帳探索時の前処理演算量をさらに削減することが可能に  
10      なるといった作用が得られる。ただしここで、音声符号化装置側で用いる拡散パターンを長さ 1 で打ち切る場合は、拡散パターンを用いない音声符号化装置に相当する (音声復号化装置には拡散パターンが適用されている)。

- また、本実施の形態では、拡散パターン格納部が、チャンネルあたり 1 種類ずつの拡散パターンを登録している場合を説明したが、チャンネルあたり 2 種類以上の  
15      拡散パターンが登録されており、それら拡散パターンを選択して用いることを特徴とするパルス拡散符号帳を確率的符号帳部に用いる音声符号化装置／音声復号化装置においても本実施の形態を適用することが可能であり、その場合にも同様の作用・効果を得ることができる。

- また、本実施の形態では、代数的符号帳部が 3 個の非零要素を含むベクトル  
20      を出力するパルス拡散符号帳を用いた場合について実施の形態を説明したが、代数的符号帳部が出力するベクトル中の非零要素数が  $M$  個 ( $M \geq 1$ ) の場合においても本実施の形態を適用することが可能であり、その場合にも同様の作用・効果を得ることができる。

- また、本実施の形態では、音声符号化装置側では、音声復号化装置側で用い  
25      る拡散パターンを半分の長さで打ち切った拡散パターンを用いる場合について説明したが、音声符号化装置側では、音声復号化装置側で用いる拡散パターンを長さ  $N$  ( $N \geq 1$ ) で打ち切り、さらに、打ち切り後の拡散パターンを  $M$  ( $M \geq 1$ )

サンプルおきにゼロに置き換えることも可能であり、その場合には、符号探索演算量をさらに低減することが可能になる。

- このように本実施の形態によれば、パルス拡散符号帳を確率的符号帳部に用
- 
- いるCELP方式の音声符号化装置と復号化装置、及び音声符号化復号化シス
- 5 テムにおいて、学習によって獲得された確率的音源ターゲット中に頻繁に含まれる固定波形を拡散パターンとして登録しておき、当該拡散パターンをパルスベクトルに重畳する（反映させる）ことで、確率的音源ターゲットにより近い確率的音源ベクトルを利用することができるため、復号化側で合成音声の品質向上を実現でき、さらには、符号化側で、パルス拡散符号帳を確率的符号帳部に用
- 10 いる場合に問題となることがある確率的符号帳探索の演算量を、従来よりも低く抑えることが可能となるという有利な効果が得られる。

なお、少数個の非零要素からなるパルスベクトルを生成する符号帳として、マルチパルス符号帳やレギュラーパルス符号帳など、その他の符号帳を用いた場合にも同様の作用・効果を得ることができる。

- 15 上記実施の形態1～3に係る音声符号化／復号化は、音声符号化装置／音声復号化装置として説明しているが、これらの音声符号化／復号化をソフトウェアとして構成しても良い。例えば、上記音声符号化／復号化のプログラムをROMに格納し、そのプログラムにしたがってCPUの指示により動作させるように構成しても良い。また、プログラム、適応符号帳、及び確率的符号帳（パ
- 20 ルス拡散符号帳）をコンピュータで読み取り可能な記憶媒体に格納し、この記憶媒体のプログラム、適応符号帳、及び確率的符号帳（パルス拡散符号帳）をコンピュータのRAMに記録して、プログラムにしたがって動作させるようにしても良い。このような場合においても、上記実施の形態1～3と同様の作用、効果を呈する。さらに、実施の形態1～3におけるプログラムを通信端末でダ
- 25 ウンロードし、その通信端末でプログラムを動作させるようにしても良い。

なお、上記実施の形態1～3については、個々に実施しても良く、組み合わせて実施しても良い。

本明細書は、1999年8月23日出願の特願平11-235050号、1999年8月24日出願の特願平11-236728号、及び1999年9月2日出願の特願平11-248363号に基づく。これらの内容はすべてここに含めておく。

5

#### 産業上の利用可能性

本発明は、デジタル通信システムにおける基地局装置や通信端末装置に適用することができる。

## 請求の範囲

1. 適応符号帳及び確率的符号帳に格納された適応音源及び確率的音源に対して、入力音声から求めたL P C係数を用いてフィルタリングすることにより、  
合成音を得るL P C合成手段と、前記適応音源及び前記確率的音源のゲインを  
5 求め、さらに前記ゲインを用いて得られる前記入力音声と前記合成音との間の  
符号化歪みを用いて適応音源及び確率的音源の符号を探索するゲイン演算手  
段と、求められた符号に対応する適応音源及び確率的音源を用いてゲインの予  
測符号化を行うパラメータ符号化手段と、を具備し、  
前記パラメータ符号化手段は、以前のサブフレームの状態に応じて前記予測  
10 符号化に用いる予測係数を調整する予測係数調整手段を備える音声符号化装  
置。  
2. 前記予測係数調整手段は、以前のサブフレームの状態が極端に大きな値又  
は極端に小さな値である場合に、その影響を小さくするように前記予測係数を  
調整する請求項1記載の音声符号化装置。  
15 3. 前記パラメータ符号化手段は、適応音源のゲインのベクトル及び確率的音  
源のゲインのベクトル、並びに予測係数を調整する係数を含む符号帳を有する  
ことを特徴とする請求項1記載の音声符号化装置。  
4. 予測符号化において、状態と予測係数との間の積和を求めるときに、その  
状態に対応した予測係数調整係数を乗じる請求項3記載の音声符号化装置。  
20 5. 前記適応音源及び前記確率的音源、並びに予測係数調整係数を状態毎に対  
応させて格納する格納手段を具備する請求項1記載の音声符号化装置。  
6. 前記格納手段に格納された前記適応音源及び前記確率的音源の状態を更新  
する際に、前記予測係数調整係数も更新する請求項5記載のベクトル量子化装  
置。  
25 7. 適応符号帳及び確率的符号帳に格納された適応音源及び確率的音源に対し  
て、入力音声から求めたL P C係数を用いてフィルタリングすることにより、  
合成音を得るL P C合成手段と、前記適応音源及び前記確率的音源のゲインを



求めるゲイン演算手段と、前記入力音声と前記合成音との間の符号化歪みを用いて求められた適応音源及び確率的音源、並びに前記ゲインのベクトル量子化を行うパラメータ符号化手段と、を具備し、1つのフレームを複数のサブフレームに分解して符号化を行うC E L P型音声符号化装置であって、

- 5      最初のサブフレームの適応符号帳探索の前に、フレームを構成する複数のサブフレームのピッチ分析を行って相関値を求め、前記相関値を用いて最もピッチ周期に近似する値を算出するピッチ分析手段を備える音声符号化装置。

8. 前記ピッチ分析手段において得られた相関値及び最もピッチ周期に近似する値に基づいて複数のサブフレームのラグの探索範囲を決定する探索範囲設

- 10     定手段を具備する請求項7記載の音声符号化装置。

9. 探索範囲設定手段は、前記ピッチ分析手段において得られた相関値及び最もピッチ周期に近似する値を用いて探索範囲の中心となる仮ピッチを求める請求項8記載の音声符号化装置。

- 15     10. 探索範囲設定手段は、仮ピッチの周りの指定の範囲にラグの探索区間を設定する請求項9記載の音声符号化装置。

11. 探索範囲設定手段は、ラグが短い候補を少なくしてラグの探索区間を設定する請求項8記載の音声符号化装置。

12. 探索範囲設定手段は、適応符号帳探索の際に、設定された範囲でラグの探索を行う請求項8記載の音声符号化装置。

- 20     13. 音声符号化プログラム；過去に合成した音源信号が格納された適応符号帳；複数の音源ベクトルを格納した確率的符号帳；を格納し、コンピュータにより読み取り可能な記録媒体であって、前記音声符号化プログラムは、

前記適応符号帳及び前記確率的符号帳に格納された適応音源及び確率的音源に対して、入力音声から求めたL P C係数を用いてフィルタリングすること

- 25     により、合成音を得る手順と、

前記適応音源及び前記確率的音源のゲインを求める手順と、

前記入力音声と前記合成音との間の符号化歪みを用いて求められた適応音

源及び確率的音源、並びに前記ゲインのベクトル量子化を行う手順と、  
を含み、

ベクトル量子化を行う手順において、複数の量子化対象ベクトルと、予測符

号化に用いる予測係数との間の符号化歪みに基づいて量子化対象ベクトルを

- 5 求める手順と、以前のサブフレームの状態に応じて前記予測係数を調整する手順と、を含む。

1 4. 音声符号化プログラム；過去に合成した音源信号が格納された適応符号帳；複数の音源ベクトルを格納した確率的符号帳；を格納し、コンピュータにより読み取り可能な記録媒体であって、前記音声符号化プログラムは、

- 10 前記適応符号帳及び前記確率的符号帳に格納された適応音源及び確率的音源に対して、入力音声から求めたL P C係数を用いてフィルタリングすることにより、合成音を得る手順と、

前記適応音源及び前記確率的音源のゲインを求める手順と、

前記入力音声と前記合成音との間の符号化歪みを用いて求められた適応音

- 15 源及び確率的音源、並びに前記ゲインのベクトル量子化を行う手順と、

最初のサブフレームの適応符号帳探索の前に、フレームを構成する複数のサブフレームのピッチ分析を行って相関値を求め、前記相関値を用いて最もピッチ周期に近似する値を算出する手順と、  
を含む。

## 補正書の請求の範囲

[2000年12月22日(22.12.00)国際事務局受理:出願当初の請求の範囲6は補正された;新しい請求の範囲15-38が加えられた;他の請求の範囲は変更なし。(6頁)]

1. 適応符号帳及び確率的符号帳に格納された適応音源及び確率的音源に対して、入力音声から求めたLPC係数を用いてフィルタリングすることにより、合成音を得るLPC合成手段と、前記適応音源及び前記確率的音源のゲインを  
5 求め、さらに前記ゲインを用いて得られる前記入力音声と前記合成音との間の符号化歪みを用いて適応音源及び確率的音源の符号を探索するゲイン演算手段と、求められた符号に対応する適応音源及び確率的音源を用いてゲインの予測符号化を行うパラメータ符号化手段と、を具備し、

前記パラメータ符号化手段は、以前のサブフレームの状態に応じて前記予測  
10 符号化に用いる予測係数を調整する予測係数調整手段を備える音声符号化装置。

2. 前記予測係数調整手段は、以前のサブフレームの状態が極端に大きな値又は極端に小さな値である場合に、その影響を小さくするように前記予測係数を調整する請求項1記載の音声符号化装置。

- 15 3. 前記パラメータ符号化手段は、適応音源のゲインのベクトル及び確率的音源のゲインのベクトル、並びに予測係数を調整する係数を含む符号帳を有することを特徴とする請求項1記載の音声符号化装置。

4. 予測符号化において、状態と予測係数との間の積和を求めるときに、その状態に対応した予測係数調整係数を乗じる請求項3記載の音声符号化装置。

- 20 5. 前記適応音源及び前記確率的音源、並びに予測係数調整係数を状態毎に対応させて格納する格納手段を具備する請求項1記載の音声符号化装置。

6. (補正後)前記格納手段に格納された前記適応音源及び前記確率的音源の状態を更新する際に、前記予測係数調整係数も更新する請求項5記載の音声符号化装置。

- 25 7. 適応符号帳及び確率的符号帳に格納された適応音源及び確率的音源に対して、入力音声から求めたLPC係数を用いてフィルタリングすることにより、合成音を得るLPC合成手段と、前記適応音源及び前記確率的音源のゲインを

求めるゲイン演算手段と、前記入力音声と前記合成音との間の符号化歪みを用いて求められた適応音源及び確率的音源、並びに前記ゲインのベクトル量子化を行うパラメータ符号化手段と、を具備し、1つのフレームを複数のサブフレームに分解して符号化を行うC E L P型音声符号化装置であって、

- 5      最初のサブフレームの適応符号帳探索の前に、フレームを構成する複数のサブフレームのピッチ分析を行って相関値を求め、前記相関値を用いて最もピッチ周期に近似する値を算出するピッチ分析手段を備える音声符号化装置。
8. 前記ピッチ分析手段において得られた相関値及び最もピッチ周期に近似する値に基づいて複数のサブフレームのラグの探索範囲を決定する探索範囲設定手段を具備する請求項7記載の音声符号化装置。
- 10     9. 探索範囲設定手段は、前記ピッチ分析手段において得られた相関値及び最もピッチ周期に近似する値を用いて探索範囲の中心となる仮ピッチを求める請求項8記載の音声符号化装置。
- 15     10. 探索範囲設定手段は、仮ピッチの周りの指定の範囲にラグの探索区間を設定する請求項9記載の音声符号化装置。
- 1     11. 探索範囲設定手段は、ラグが短い候補を少なくしてラグの探索区間を設定する請求項8記載の音声符号化装置。
- 1     12. 探索範囲設定手段は、適応符号帳探索の際に、設定された範囲でラグの探索を行う請求項8記載の音声符号化装置。
- 20     13. 音声符号化プログラム；過去に合成した音源信号が格納された適応符号帳；複数の音源ベクトルを格納した確率的符号帳；を格納し、コンピュータにより読み取り可能な記録媒体であって、前記音声符号化プログラムは、  
前記適応符号帳及び前記確率的符号帳に格納された適応音源及び確率的音源に対して、入力音声から求めたL P C係数を用いてフィルタリングすること  
25     により、合成音を得る手順と、  
前記適応音源及び前記確率的音源のゲインを求める手順と、  
前記入力音声と前記合成音との間の符号化歪みを用いて求められた適応音

源及び確率的音源、並びに前記ゲインのベクトル量子化を行う手順と、  
を含み、

ベクトル量子化を行う手順において、複数の量子化対象ベクトルと、予測符号化に用いる予測係数との間の符号化歪みに基づいて量子化対象ベクトルを  
5 求める手順と、以前のサブフレームの状態に応じて前記予測係数を調整する手順と、を含む。

1 4. 音声符号化プログラム；過去に合成した音源信号が格納された適応符号帳；複数の音源ベクトルを格納した確率的符号帳；を格納し、コンピュータにより読み取り可能な記録媒体であって、前記音声符号化プログラムは、

10 前記適応符号帳及び前記確率的符号帳に格納された適応音源及び確率的音源に対して、入力音声から求めたL P C係数を用いてフィルタリングすることにより、合成音を得る手順と、

前記適応音源及び前記確率的音源のゲインを求める手順と、

前記入力音声と前記合成音との間の符号化歪みを用いて求められた適応音  
15 源及び確率的音源、並びに前記ゲインのベクトル量子化を行う手順と、

最初のサブフレームの適応符号帳探索の前に、フレームを構成する複数のサブフレームのピッチ分析を行って相関値を求め、前記相関値を用いて最もピッチ周期に近似する値を算出する手順と、  
を含む。

20 1 5. (追加) 少なくとも一つの非零要素（非零要素以外の要素はゼロの値を持つ）を含むベクトルと、拡散パターンと呼ばれる固定波形とを重畳してベクトルを生成するパルス拡散符号帳を備え、前記パルス拡散符号帳が、音声復号化装置側のパルス拡散符号帳の構成と異なる構成を有する音声符号化装置。

1 6. (追加) パルス拡散符号帳の構成部位である拡散パターン格納部が、音声  
25 復号化装置側の拡散パターン格納部が格納している拡散パターンと異なる拡散パターンを格納している請求項 1 5 記載の音声符号化装置。

1 7. (追加) 拡散パターン格納部が、音声復号化装置側の拡散パターン格納部が

格納している拡散パターンを簡素化して選られる得られる拡散パターンを格納している請求項 1 6 記載の音声符号化装置。

1 8. (追加) 拡散パターン格納部が、音声復号化装置側の拡散パターン格納部が格納している拡散パターの構成要素を、所定の間隔ごとにゼロに置き換えて得られる拡散パターンを格納している請求項 1 6 記載の音声符号化装置。

1 9. (追加) 拡散パターン格納部が、音声復号化装置側の拡散パターン格納部が格納している拡散パターの構成要素を、Nサンプル(Nは自然数)ごとにゼロに置き換えて得られる拡散パターンを格納している請求項 1 6 記載の音声符号化装置。

10 2 0. (追加) 拡散パターン格納部が、音声復号化装置側の拡散パターン格納部が格納している拡散パターの構成要素を、1サンプルごとにゼロに置き換えて得られる拡散パターンを格納している請求項 1 9 記載の音声符号化装置。

2 1. (追加) 拡散パターン格納部が、音声復号化装置側の拡散パターン格納部が格納している拡散パターの構成要素を、適当な長さで打ち切って得られる拡散パターンを格納している請求項 1 6 記載の音声符号化装置。

2 2. (追加) 拡散パターン格納部が、音声復号化装置側の拡散パターン格納部が格納している拡散パターの構成要素を、Nサンプル(Nは自然数)の長さで打ち切って得られる拡散パターンを格納していることを特徴とする請求項 1 6 記載の音声符号化装置。

20 2 3. (追加) 拡散パターン格納部が、音声復号化装置側の拡散パターン格納部が格納している拡散パターの構成要素を、半分の長さに打ち切って得られる拡散パターンを格納している請求項 1 6 記載の音声符号化装置。

2 4. (追加) 請求項 1 5 音声符号化装置で生成された音声符号を有する音声信号を復号化する音声復号化装置。

25 2 5. (追加) 請求項 1 5 音声符号化装置を実現するソフトウェアプログラムを記述した信号処理用プロセッサ。

2 6. (追加) 請求項 2 4 記載の音声復号化装置を実現するソフトウェアプロ

グラムを記述した信号処理用プロセッサ。

27. (追加) 音声符号化装置側が有するパルス拡散符号帳の構成と、音声復号化装置側が有するパルス拡散符号帳の構成とが異なる音声符号化復号化システム。

5 28. (追加) 音声符号化装置側が有するパルス拡散符号帳の構成と、音声符号化装置側が有するパルス拡散符号帳の構成との違いが、それぞれのパルス拡散符号帳に備えられた拡散パタンの形状である請求項27記載の音声符号化復号化システム。

29. (追加) 音声符号化装置側の拡散パタンの形状が、音声復号化装置側の  
10 拡散パタンの形状を簡素化したのものである請求項28記載の音声符号化復号化システム。

30. (追加) 音声符号化装置側の拡散パタンの形状が、音声復号化装置側の拡散パタンの構成要素を、適当な間隔ごとにゼロに置き換えて得られる形状である請求項27記載の音声符号化復号化システム。

15 31. (追加) 音声符号化装置側の拡散パタンの形状が、音声復号化装置側の拡散パタンの構成要素を、Nサンプル(Nは自然数)ごとにゼロに置き換えて得られる形状である請求項27記載の音声符号化復号化システム。

32. (追加) 音声符号化装置側の拡散パタンの形状が、音声復号化装置側の拡散パタンの構成要素を、1サンプルごとにゼロに置き換えて得られる形状で  
20 ある請求項31記載の音声符号化復号化システム。

33. (追加) 音声符号化装置側の拡散パタンの形状が、音声復号化装置側の拡散パタンの構成要素を、適当な長さで打ち切って得られる形状である請求項27記載の音声符号化復号化システム。

34. (追加) 音声符号化装置側の拡散パタンの形状が、音声符号化装置側の  
25 拡散パタンの構成要素を、Nサンプル(Nは自然数)の長さで打ち切って得られる形状である請求項27記載の音声符号化復号化システム。

35. (追加) 音声符号化装置側の拡散パタンの形状が、音声復号化装置側の

拡散パタンの構成要素を、半分の長さに打ち切って得られる形状である請求項 2 7 記載の音声符号化復号化システム。

---

- 3 6. (追加) 請求項 2 5 記載の信号処理用プロセッサを備える通信用基地局。
- 5 3 7. (追加) 請求項 2 5 記載の信号処理用プロセッサを備える通信用端末。
- 3 8. (追加) 請求項 3 6 記載の通信用基地局及び請求項 3 7 記載の通信端末を無線ネットワークでつないだ無線通信システム。



1 / 14

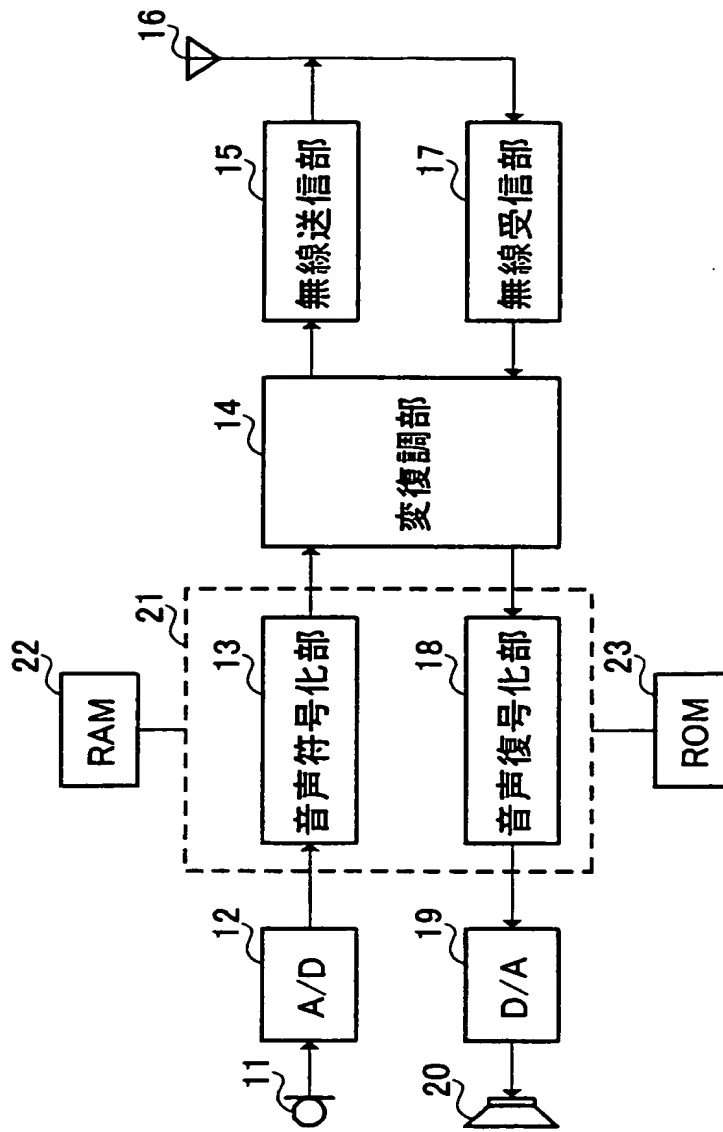


図 1

---

***THIS PAGE BLANK (USPTO)***

2/14

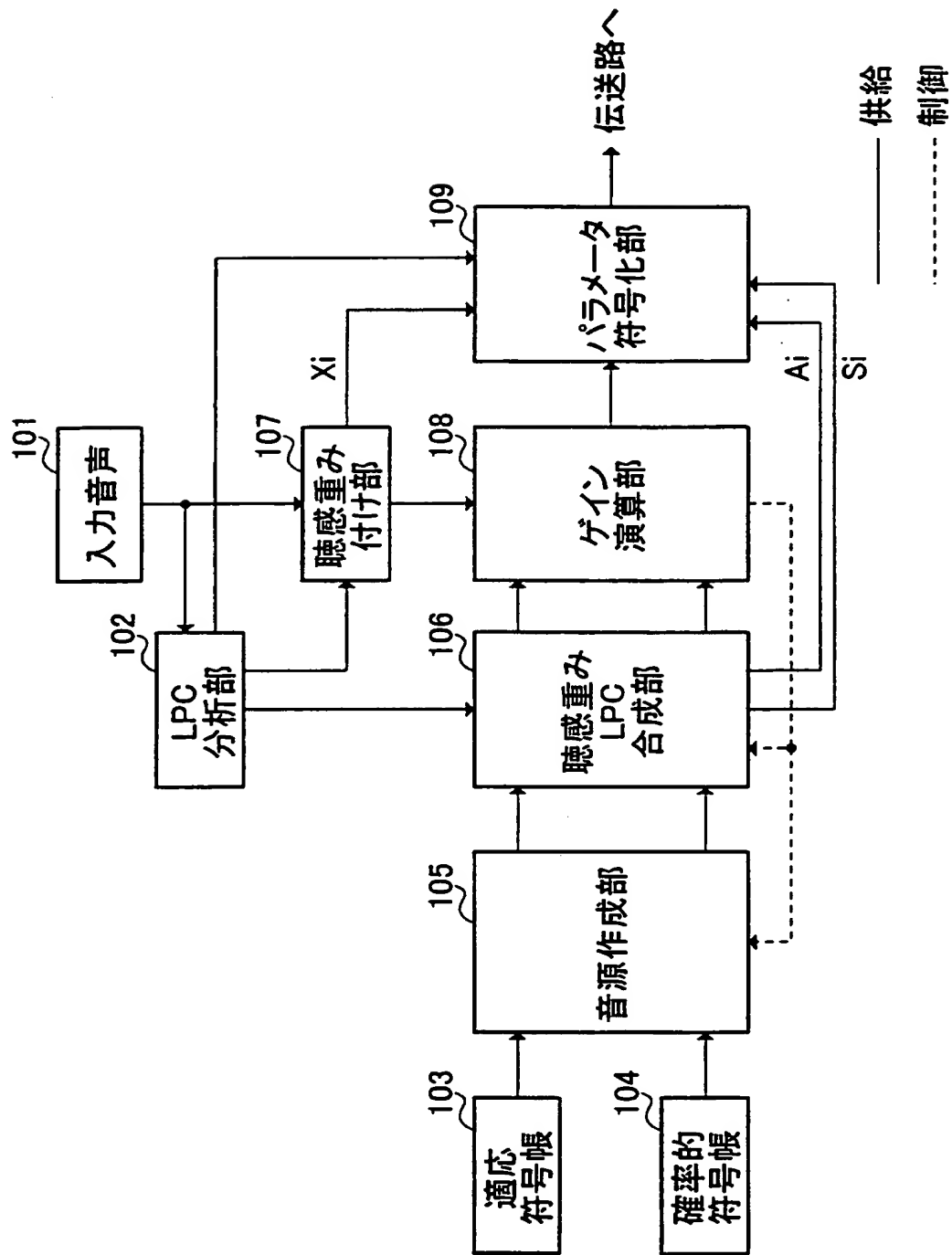
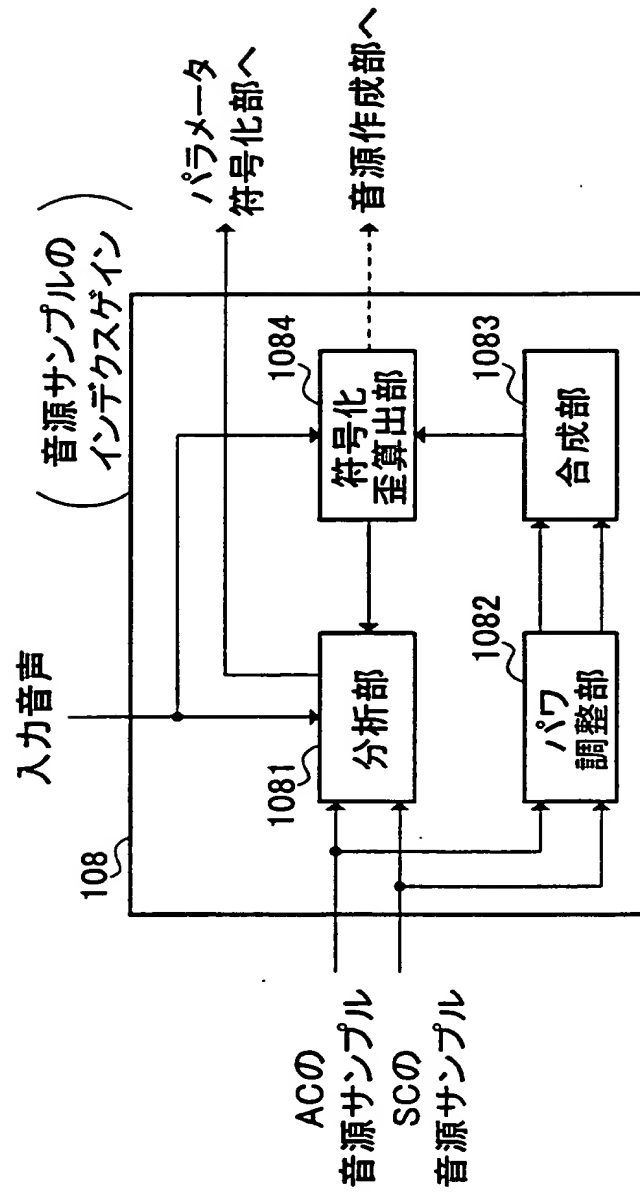


図 2

**THIS PAGE BLANK (USPTO)**



---

***THIS PAGE BLANK (USPTO)***

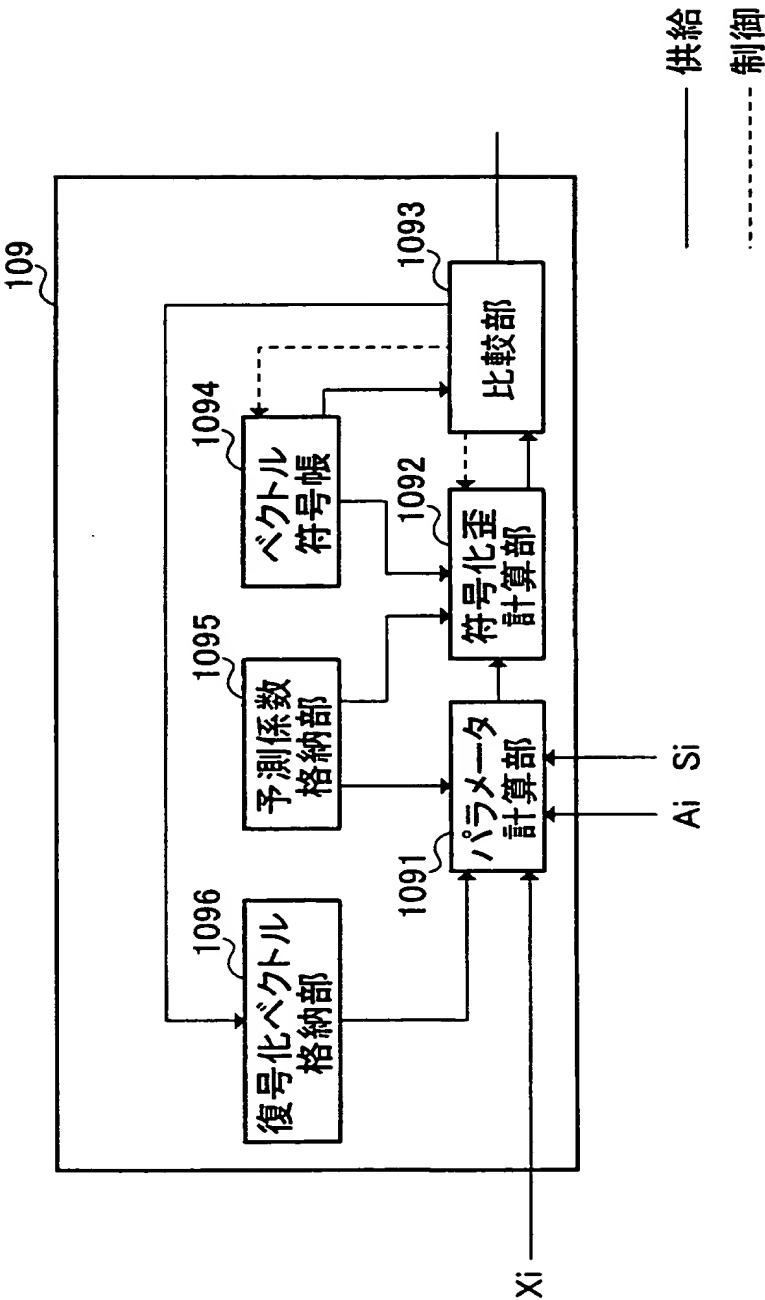


図 4

---

**THIS PAGE BLANK (USPTO)**



5/14

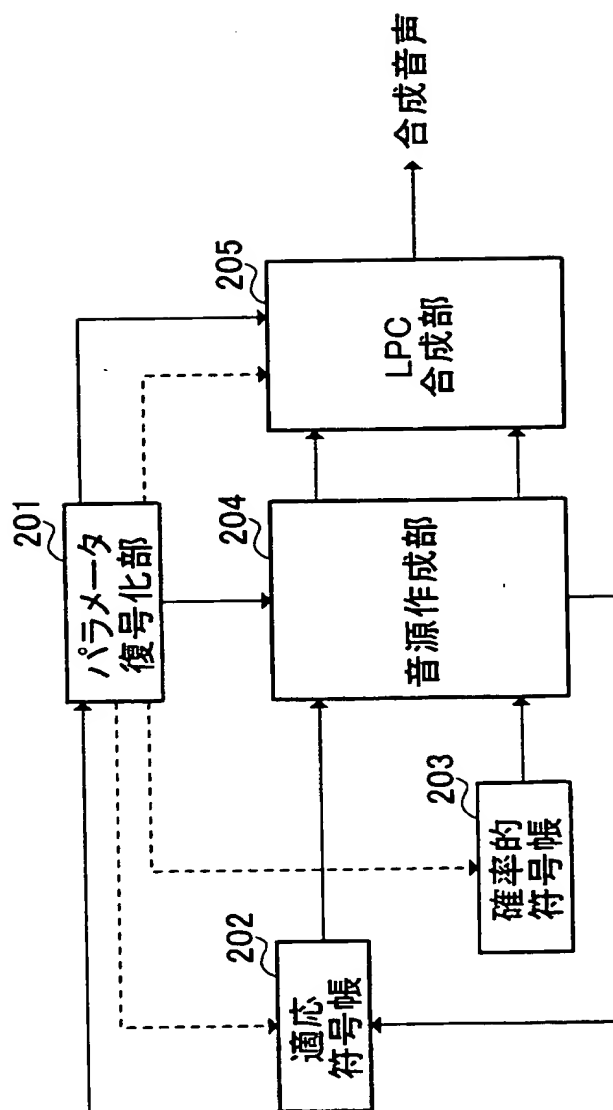


図 5

---

**THIS PAGE BLANK (USPTO)**

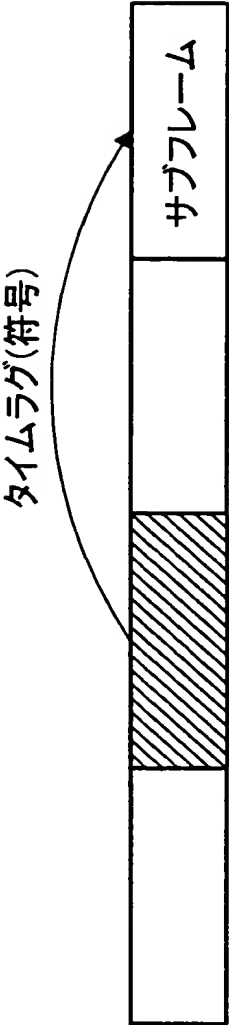


図 6

---

**THIS PAGE BLANK (USPTO)**

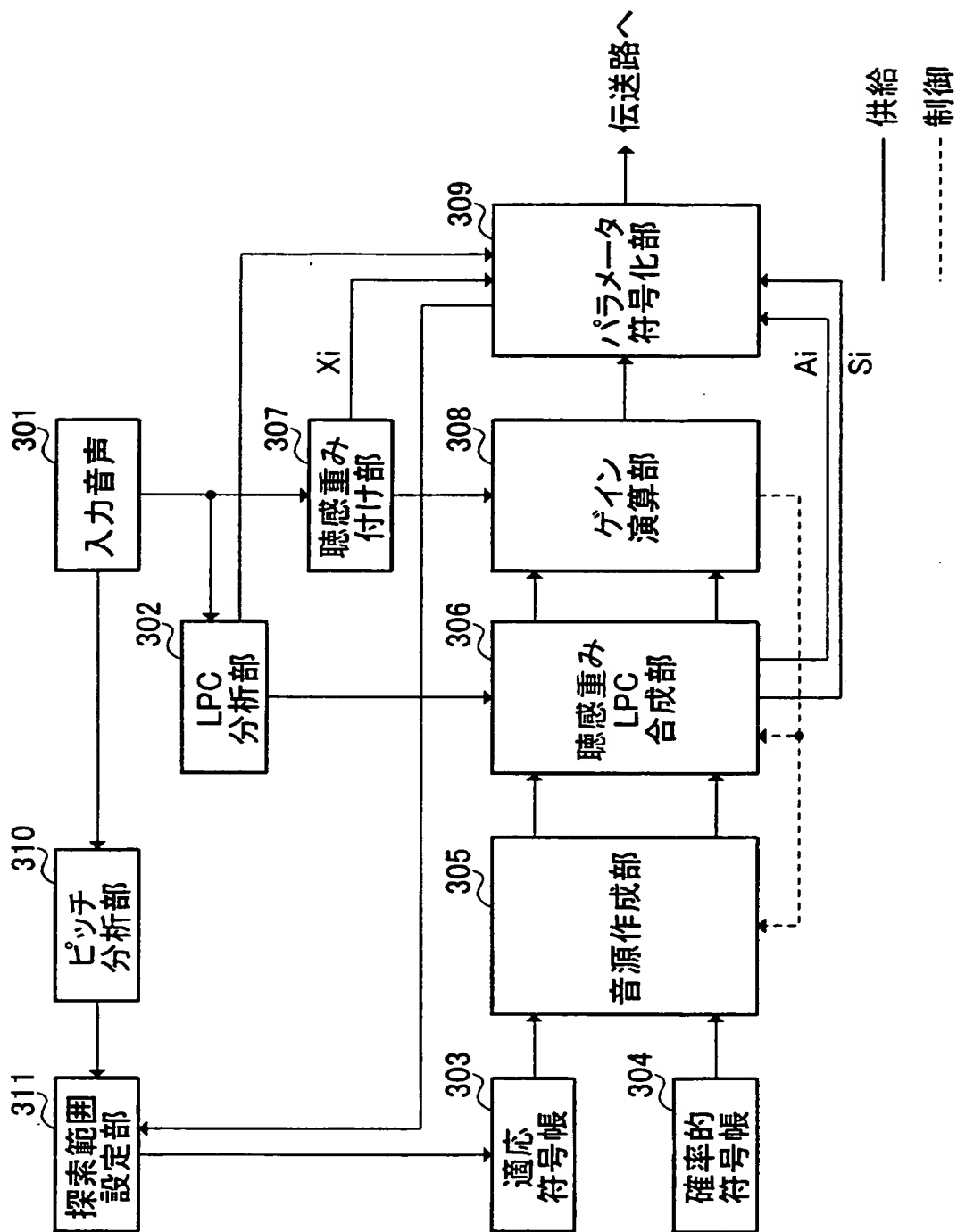


図 7

**THIS PAGE BLANK (USPTO)**

8/14

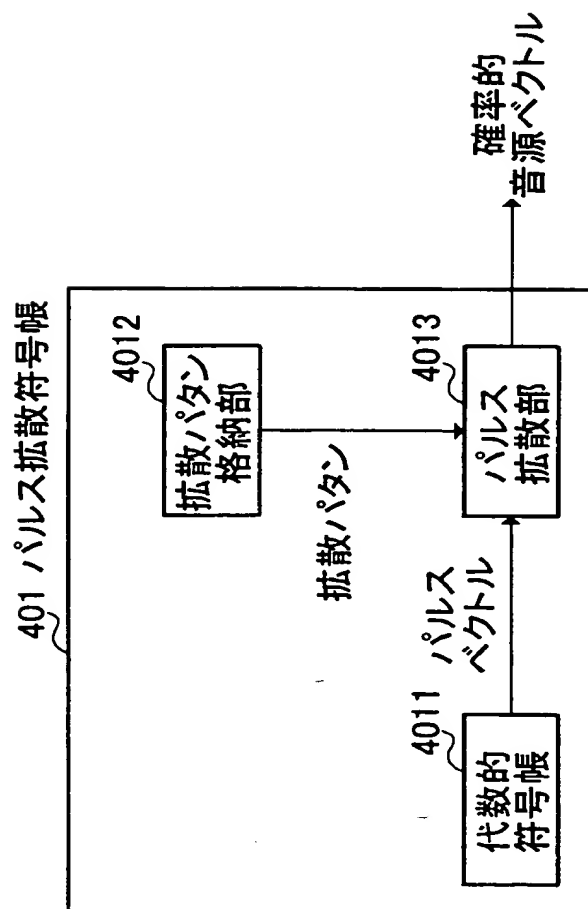


図 8

---

**THIS PAGE BLANK (USPTO)**



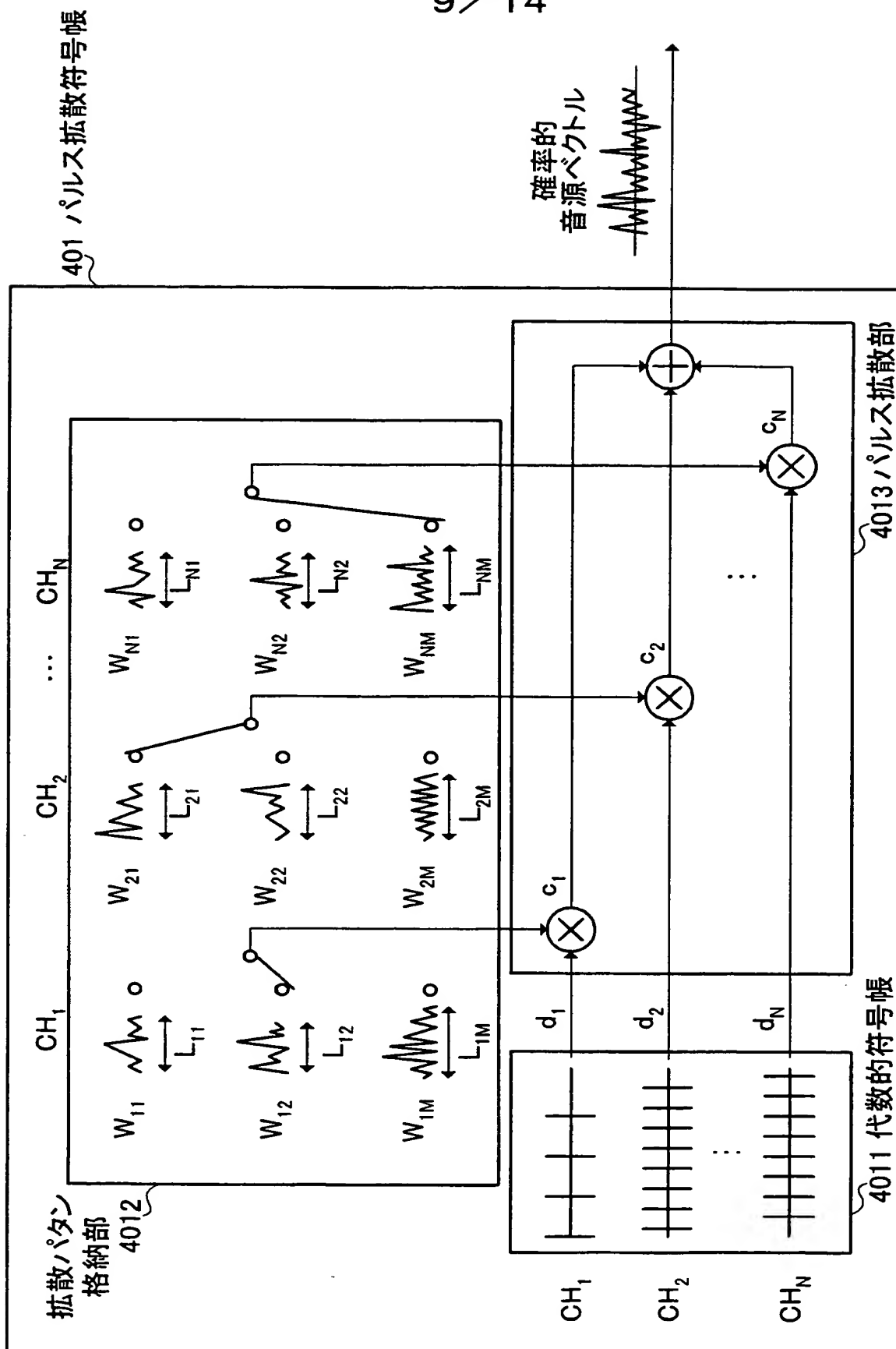


図 9

---

**THIS PAGE BLANK (USPTO)**

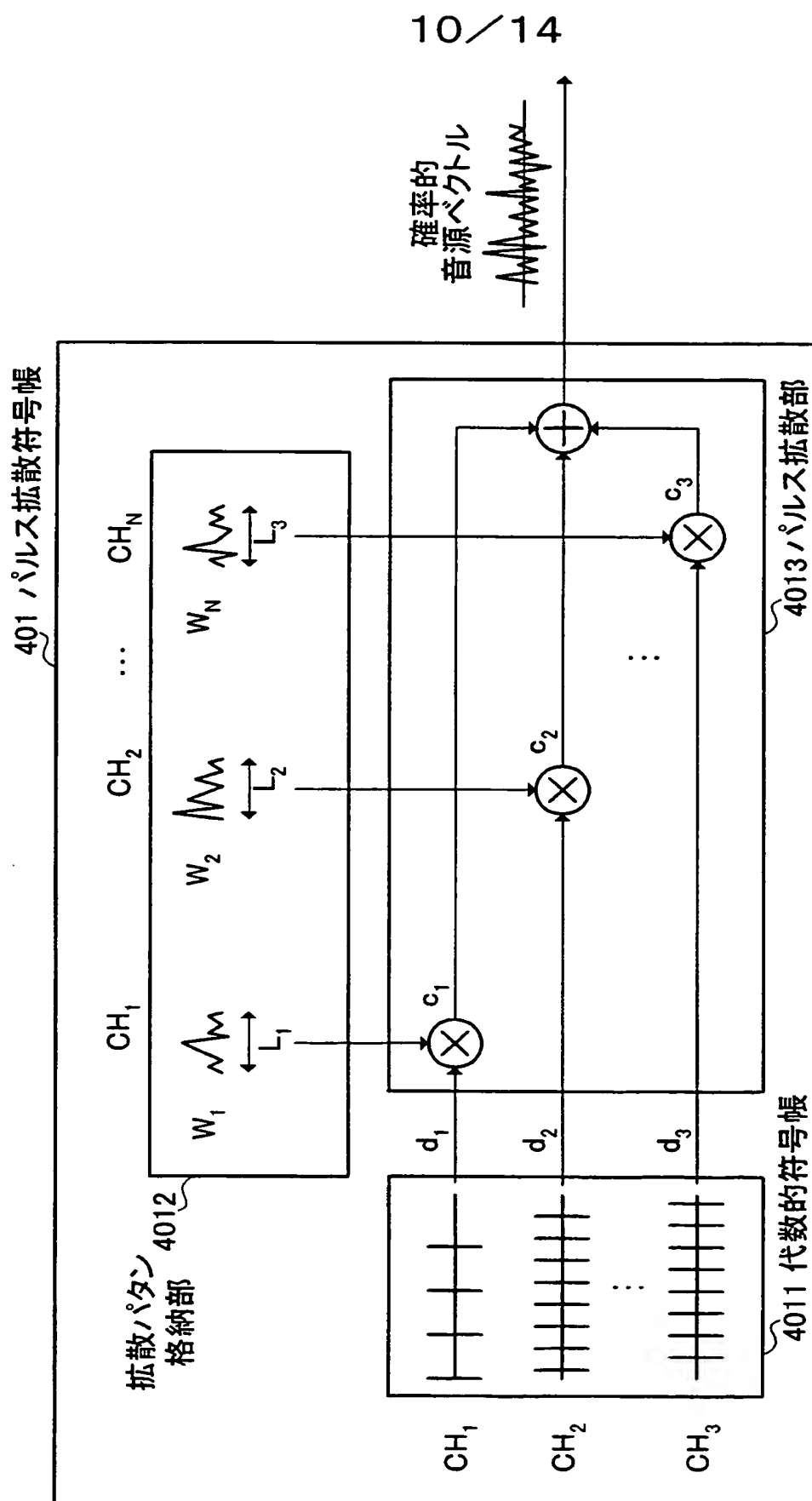


図 10

---

**THIS PAGE BLANK (USPTO)**

11 / 14

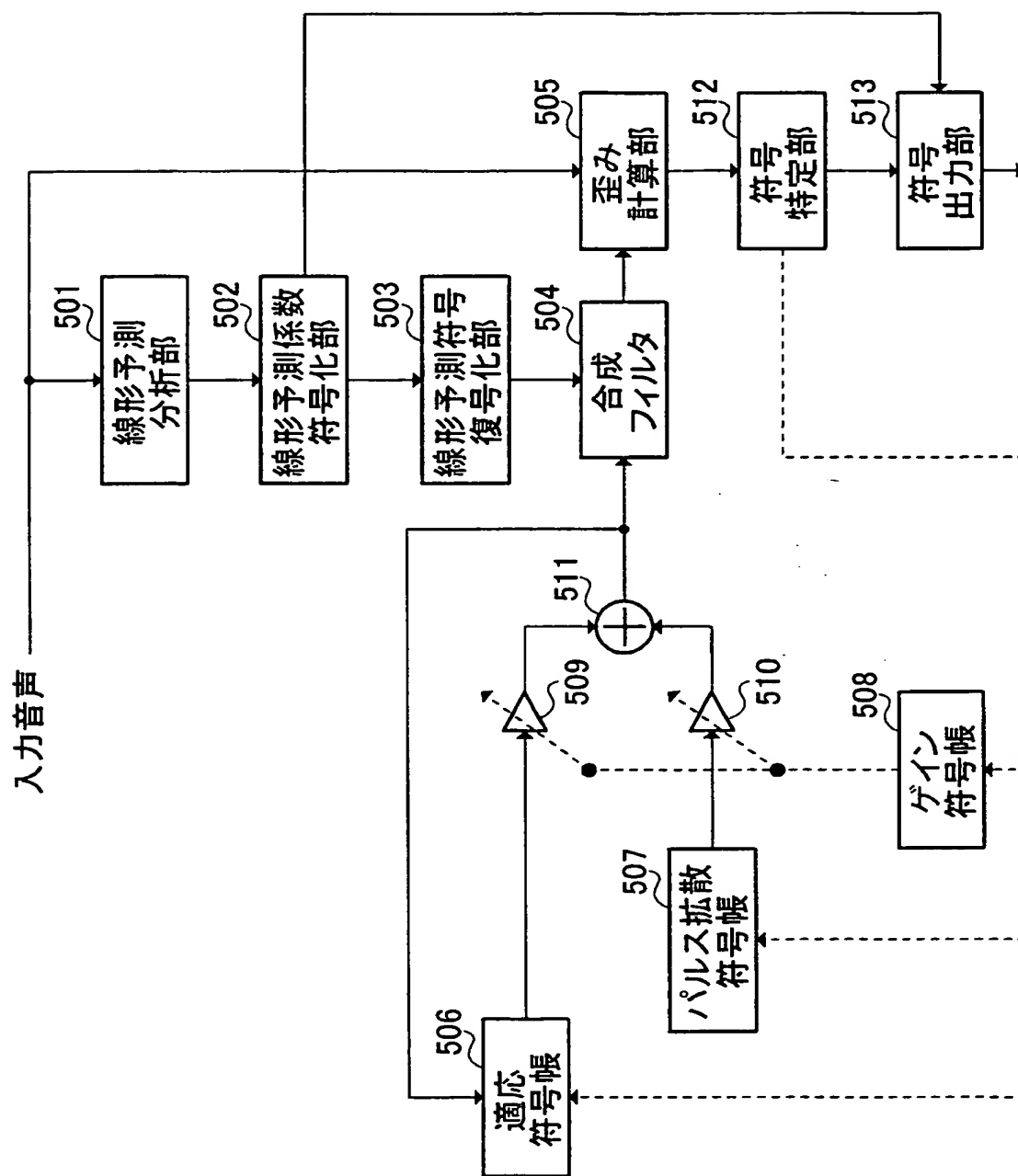


図 11

**THIS PAGE BLANK (USPTO)**

12/14

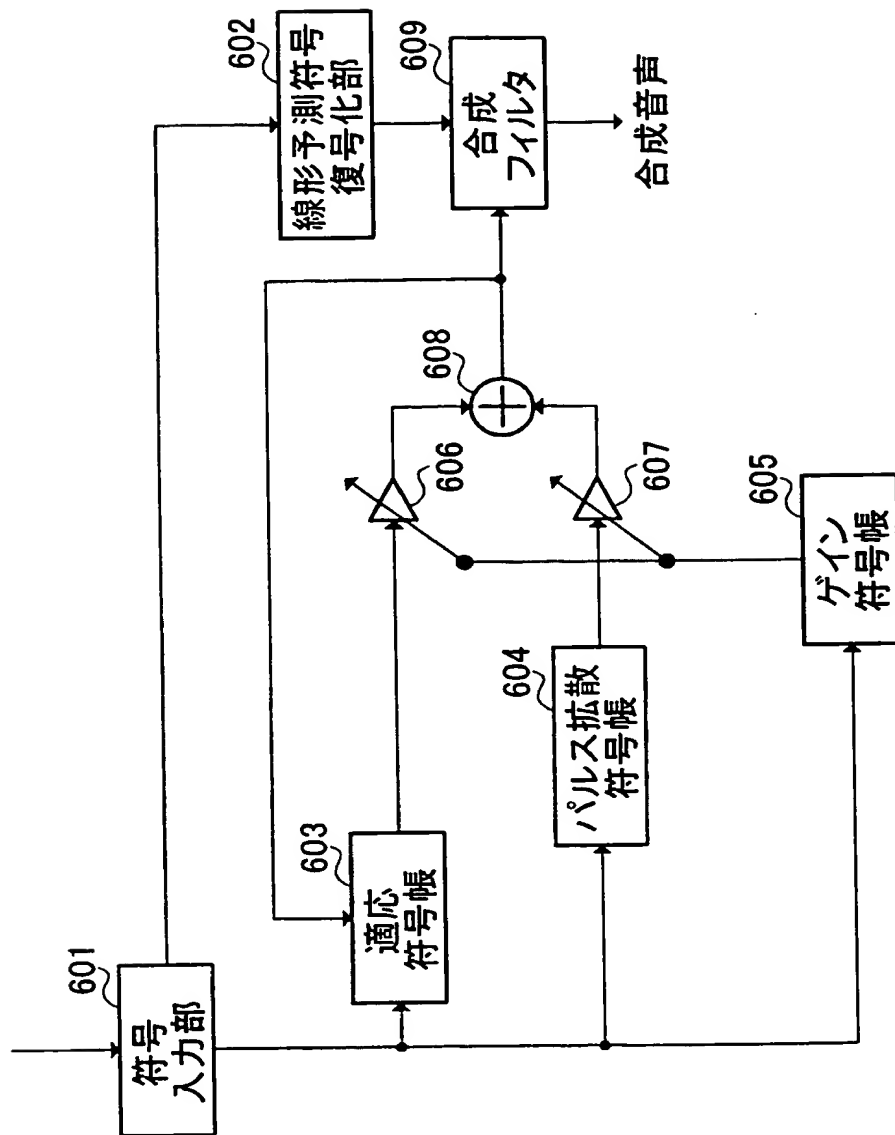


図 12

**THIS PAGE BLANK (USPTO)**



13/14

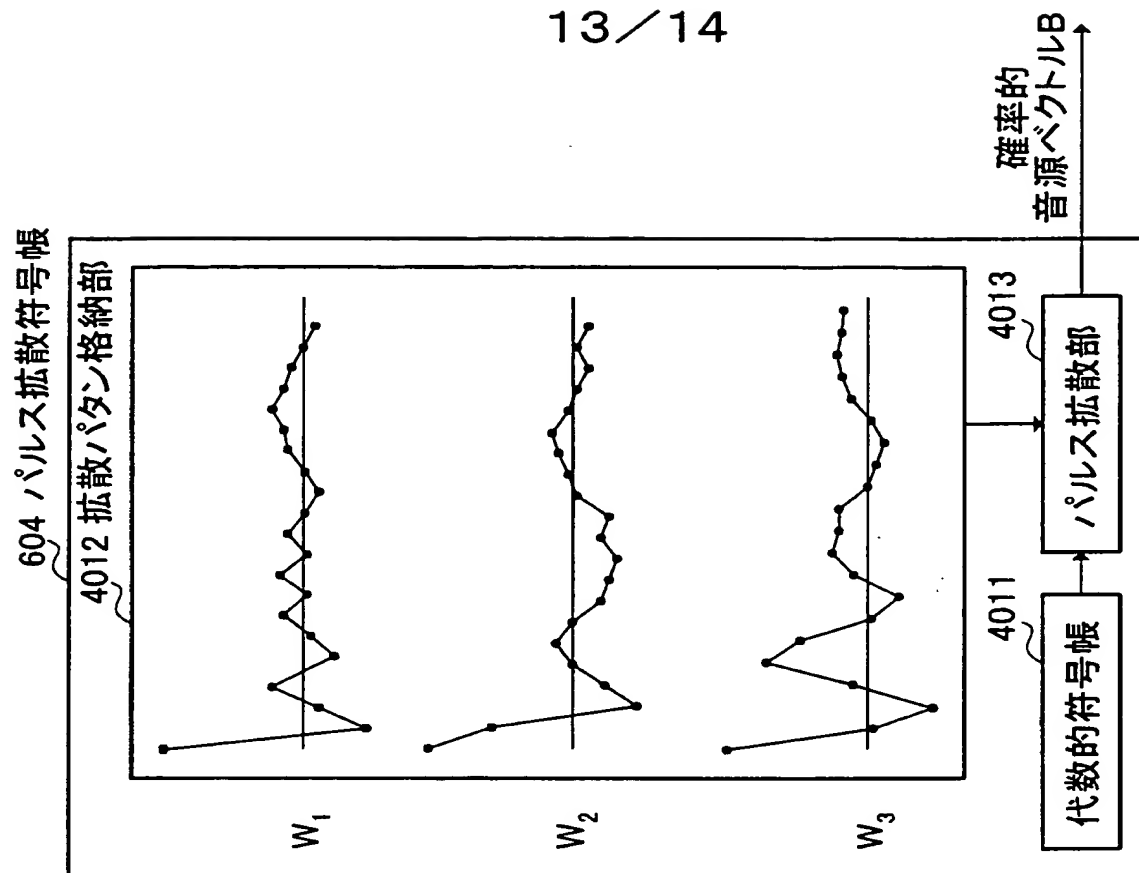


図 13B

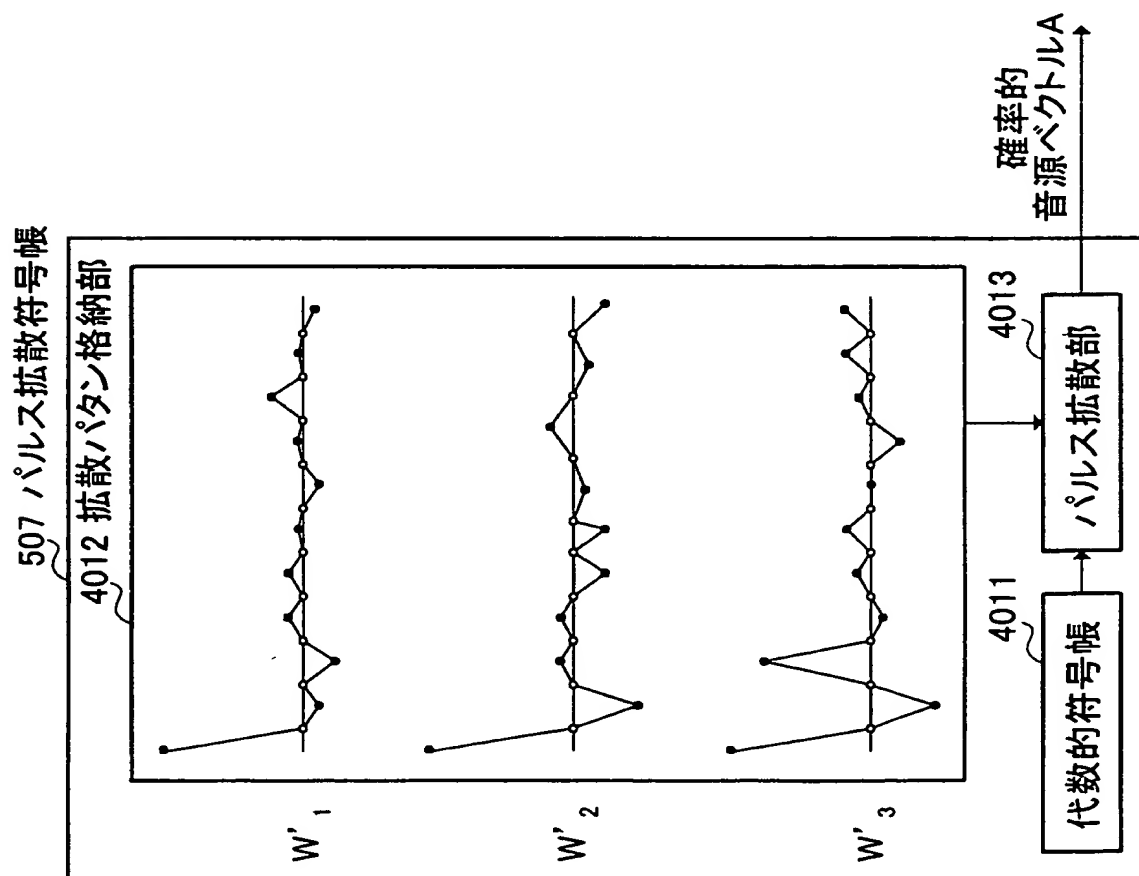


図 13A

**THIS PAGE BLANK (USPTO)**

14/14

604 パルス拡散符号帳

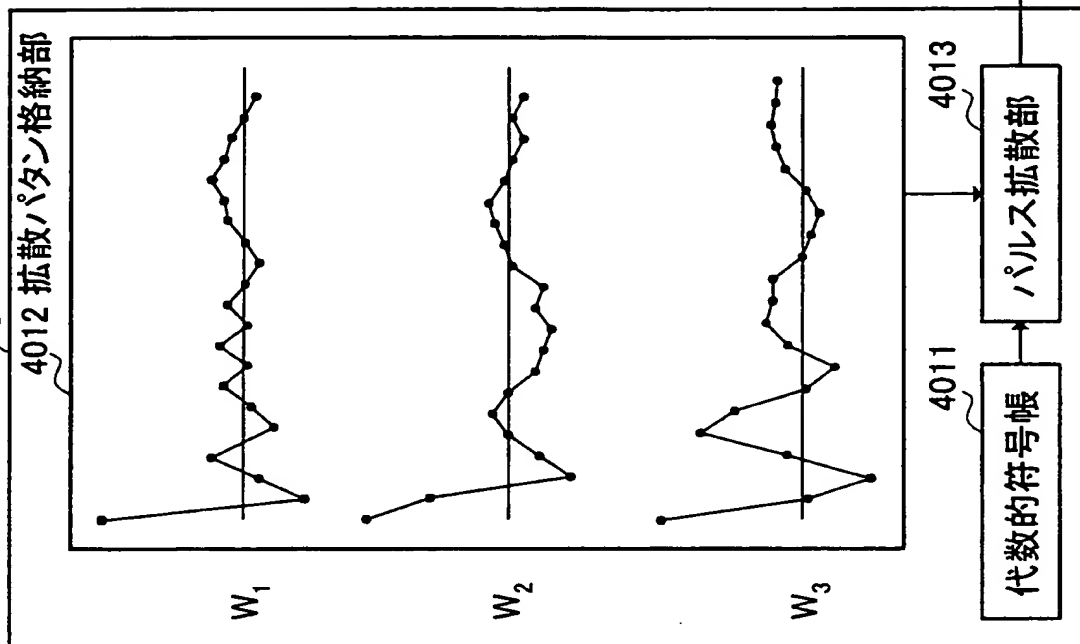


図 14B

507 パルス拡散符号帳

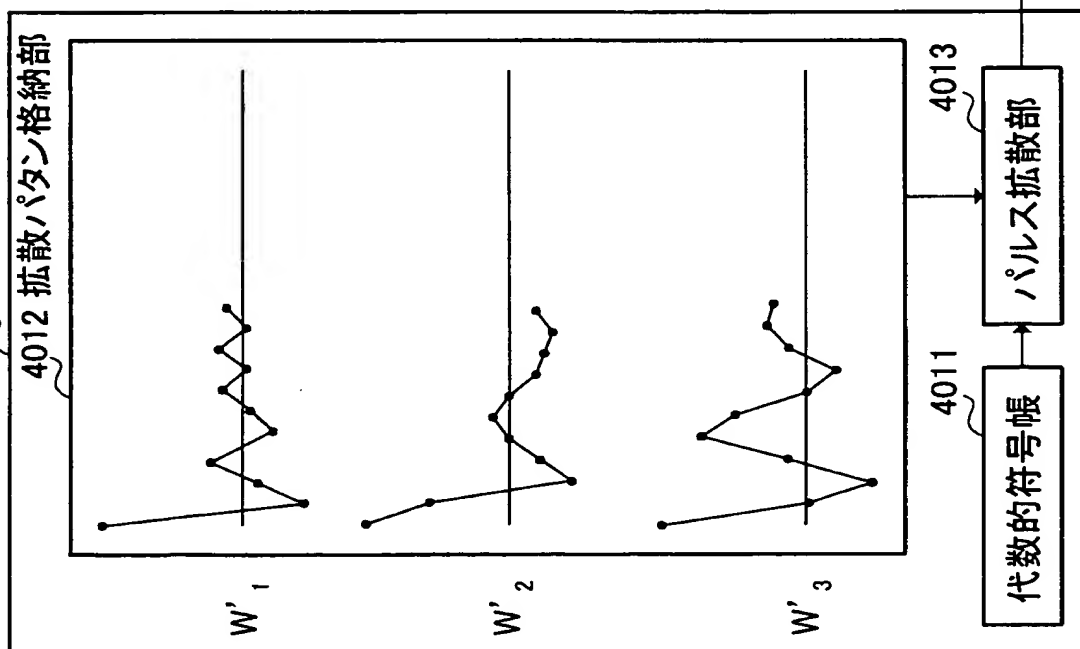


図 14A

**THIS PAGE BLANK (USPTO)**

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP00/05621

## A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl<sup>7</sup> G10L19/04  
//G10L101:12

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl<sup>7</sup> G10L19/04

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Toroku Jitsuyo Shinan Koho	1994-2000
Kokai Jitsuyo Shinan Koho	1971-2000	Jitsuyo Shinan Toroku Koho	1996-2000

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

JICST FILE (JOIS)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP, 10-233694, A (Matsushita Electric Ind. Co., Ltd.), 02 September, 1998 (02.09.98) & WO, 98/20483, A1 & AU, 4884297, A & EP, 883107, A1	1-14
A	JP, 10-282998, A (Matsushita Electric Ind. Co., Ltd.), 23 October, 1998 (23.10.98) (Family: none)	1-6, 13
A	JP, 9-152897, A (Hitachi, Ltd.), 10 June, 1997 (10.06.97) (Family: none)	7-12, 14

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&amp;" document member of the same patent family

Date of the actual completion of the international search  
06 October, 2000 (06.10.00)Date of mailing of the international search report  
17 October, 2000 (17.10.00)Name and mailing address of the ISA/  
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

**THIS PAGE BLANK (USPTO)**